

TRAITE DE COOPERATION EN MATIERE DE BREVETS

PCT

NOTIFICATION D'ELECTION

(règle 61.2 du PCT)

Expéditeur: le BUREAU INTERNATIONAL

Destinataire:

Commissioner
 US Department of Commerce
 United States Patent and Trademark
 Office, PCT
 2011 South Clark Place Room
 CP2/5C24
 Arlington, VA 22202
 ETATS-UNIS D'AMERIQUE
 en sa qualité d'office élu

Date d'expédition (jour/mois/année) 20 mars 2001 (20.03.01)	
Demande internationale no PCT/FR00/01909	Référence du dossier du déposant ou du mandataire BCT000055
Date du dépôt international (jour/mois/année) 04 juillet 2000 (04.07.00)	Date de priorité (jour/mois/année) 05 juillet 1999 (05.07.99)
Déposant CAPMAN, François etc	

1. L'office désigné est avisé de son élection qui a été faite:



dans la demande d'examen préliminaire international présentée à l'administration chargée de l'examen préliminaire international le:

22 janvier 2001 (22.01.01)



dans une déclaration visant une élection ultérieure déposée auprès du Bureau international le:

2. L'élection



a été faite



n'a pas été faite

avant l'expiration d'un délai de 19 mois à compter de la date de priorité ou, lorsque la règle 32 s'applique, dans le délai visé à la règle 32.2b).

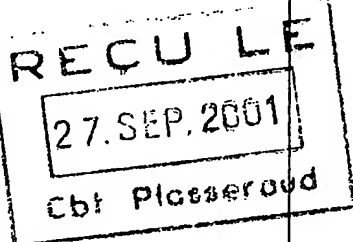
Bureau international de l'OMPI 34, chemin des Colombettes 1211 Genève 20, Suisse no de télécopieur: (41-22) 740.14.35	Fonctionnaire autorisé Kiwa Mpay no de téléphone: (41-22) 338.83.38
--	---

THIS PAGE BLANK (USPTO)

Expéditeur: L'ADMINISTRATION CHARGÉE DE
L'EXAMEN PRÉLIMINAIRE INTERNATIONAL

Destinataire:

LOISEL, Bertrand
CABINET PLASSERAUD
84, rue d'Amsterdam
75440 Paris Cedex 09
FRANCE



PCT

**NOTIFICATION DE TRANSMISSION DU
RAPPORT D'EXAMEN PRÉLIMINAIRE
INTERNATIONAL**

(règle 71.1 du PCT)

Date d'expédition
(jour/mois/année) 25.09.2001

Référence du dossier du déposant ou du mandataire
BCT000055 BLO/EB/SDV

NOTIFICATION IMPORTANTE

Demande internationale No.
PCT/FR00/01909

Date du dépôt international (jour/mois/année)
04/07/2000

Date de priorité (jour/mois/année)
05/07/1999

Déposant

MATRA NORTEL COMMUNICATIONS et al.

1. Il est notifié au déposant que l'administration chargée de l'examen préliminaire international a établi le rapport d'examen préliminaire international pour la demande internationale et le lui transmet ci-joint, accompagné, le cas échéant, de ces annexes.
2. Une copie du présent rapport et, le cas échéant, de ses annexes est transmise au Bureau international pour communication à tous les offices élus.
3. Si tel ou tel office élu l'exige, le Bureau international établira une traduction en langue anglaise du rapport (à l'exclusion des annexes de celui-ci) et la transmettra aux offices intéressés.

4. RAPPEL

Pour aborder la phase nationale auprès de chaque office élu, le déposant doit accomplir certains actes (dépôt de traduction et paiement des taxes nationales) dans le délai de 30 mois à compter de la date de priorité (ou plus tard pour ce qui concerne certains offices) (article 39.1) (voir aussi le rappel envoyé par le Bureau international dans le formulaire PCT/IB/301).

Lorsqu'une traduction de la demande internationale doit être remise à un office élu, elle doit comporter la traduction de toute annexe du rapport d'examen préliminaire international. Il appartient au déposant d'établir la traduction en question et de la remettre directement à chaque office élu intéressé.

Pour plus de précisions en ce qui concerne les délais applicables et les exigences des offices élus, voir le Volume II du Guide du déposant du PCT.

Nom et adresse postale de l'administration chargée de l'examen préliminaire international

Office européen des brevets
D-80298 Munich
Tél. +49 89 2399 - 0 Tx: 523656 epmu d
Fax: +49 89 2399 - 4465



Fonctionnaire autorisé

Benigar, M

Tél.+49 89 2399-2996



THIS PAGE BLANK (USPTO)

187 819 189

Translation

PATENT COOPERATION TREATY

PCT

#6 6-18-02
Translation

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference BCT000055	FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/FR00/01909	International filing date (day/month/year) 04 July 2000 (04.07.00)	Priority date (day/month/year) 05 July 1999 (05.07.99)
International Patent Classification (IPC) or national classification and IPC G10L 19/02		
Applicant MATRA NORTEL COMMUNICATIONS		

- This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
- This REPORT consists of a total of 6 sheets, including this cover sheet.

☒ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

These annexes consist of a total of 2 sheets.

- This report contains indications relating to the following items:

- I ☒ Basis of the report
- II ☐ Priority
- III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- IV ☐ Lack of unity of invention
- V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI ☐ Certain documents cited
- VII ☒ Certain defects in the international application
- VIII ☐ Certain observations on the international application

Date of submission of the demand. 22 January 2001 (22.01.01)	Date of completion of this report 25 September 2001 (25.09.2001)
Name and mailing address of the IPEA/EP. Facsimile No.	Authorized officer Telephone No.

THIS PAGE BLANK (USPTO)

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/FR00/01909

I. Basis of the report

1. With regard to the **elements** of the international application:*

- ☐ the international application as originally filed
- ☒ the description:
pages 1-32, as originally filed
pages _____, filed with the demand
pages _____, filed with the letter of _____
- ☒ the claims:
pages 1-10, 11, 17, 18-24, as originally filed
pages _____, as amended (together with any statement under Article 19
pages _____, filed with the demand
pages 11, 12-16, 17, 25-29, filed with the letter of 07 June 2001 (07.06.2001)
- ☒ the drawings:
pages 1/13-13/13, as originally filed
pages _____, filed with the demand
pages _____, filed with the letter of _____
- ☐ the sequence listing part of the description:
pages _____, as originally filed
pages _____, filed with the demand
pages _____, filed with the letter of _____

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language _____ which is:

- ☐ the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of the translation furnished for the purposes of international preliminary examination (under Rule 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

- ☐ contained in the international application in written form.
- ☐ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☐ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. ☐ The amendments have resulted in the cancellation of:

- ☐ the description, pages _____
- ☐ the claims, Nos. _____
- ☐ the drawings, sheets/fig _____

5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).**

*. Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rule 70.16 and 70.17).

** Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.

THIS PAGE BLANK (USPTO)

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Claims	1-29	YES
	Claims		NO
Inventive step (IS)	Claims	1-29	YES
	Claims		NO
Industrial applicability (IA)	Claims	1-29	YES
	Claims		NO

2. Citations and explanations

1. The prior art most relevant for the application is considered to be D1: AHMADI S ET AL: 'New techniques for sinusoidal coding of speech at 2400 bps' CONFERENCE RECORD OF THIRTIETH ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS (CAT. NO. 96 CB36004), PACIFIC GROVE, CA, USA, 3-6 NOV. 1996, pages 770-774 vol. 1, XP002138769 1997, Los Alamitos, CA, USA, IEEE Comput. Soc. Press, USA ISBN: 0-8186-7646-9.
2. Using the wording of Claim 1, D1 discloses a method for coding an audio signal (page 770, first 3 lines of the introduction) in which a fundamental frequency is estimated (page 770, left-hand column, last 2 lines), a spectrum is determined by transformation in the frequency domain (LPC, page 770, right-hand column, lines 16-25), and data representing spectral amplitude at multiples of the fundamental frequency are included in a digital data flow (A^{km} , Φ^{km} , page 770, right-hand column, lines 20-23, see also Chapter 3, first paragraph, page 771). In D1, the representative data A^{km} are obtained by sampling the LPC envelope at multiples of the fundamental frequency, the value of which has

THIS PAGE BLANK (USPTO

been determined by means of cepstral coefficients calculated by transforming, in the cepstral domain, the compressed spectrum (by applying a logarithm, as is common practice) of the audio signal. Moreover, the data contain a Φ_m^k phase information, which represents the phase of the spectrum at multiple frequencies of the fundamental frequency (see page 770, right-hand column, lines 4-15, Φ^{km} is associated with A_m^k , which is the amplitude of the multiple line m of the segment k).

Consequently, only the feature according to which this is an upper envelope of the audio spectrum that is transformed to obtain the cepstral coefficients is novel over D1.

3. In D1, the cepstrum is calculated in order to determine the fundamental frequency information, which can easily be read on said cepstrum, as is well known. It is therefore necessary to roughly calculate the inverse transformation of the spectrum module logarithm. If indeed only the spectrum envelope is considered, the spectral line information would disappear and the fundamental frequency calculation would no longer be possible on the resulting cepstrum. Therefore, there is no reason for a person skilled in the art to modify the teaching of D1 in order to add thereto the cepstral calculation from an upper envelope of the audio spectrum, as claimed.

This novel feature, which in the application serves other purposes, must therefore be considered to be inventive.

THIS PAGE BLANK (USPTO)

4. Claims 2 to 15 are considered to be novel and inventive for the same reasons as Claim 1 on which they are dependent.
5. The audio encoder of Claim 16, programmed to carry out a method judged to be novel and inventive, is likewise considered to be novel and inventive.
6. The features of determining phase as part of the decoding, as claimed in Claim 17 are neither known nor made obvious by D1 or by another document. Claim 17 is therefore considered to be novel and inventive.
7. Claims 18 to 25 are considered to be novel and inventive for the same reason as Claim 17 on which they are dependent.
8. The features of determining phase variation as part of the decoding, as claimed in Claim 26, are neither known nor made obvious by D1 or by another document. Claim 26 is therefore considered to be novel and inventive.
9. The audio decoder of Claim 29, which is programmed to carry out a method judged to be novel and inventive, is likewise considered to be novel and inventive.
10. All of the claims are industrially applicable.

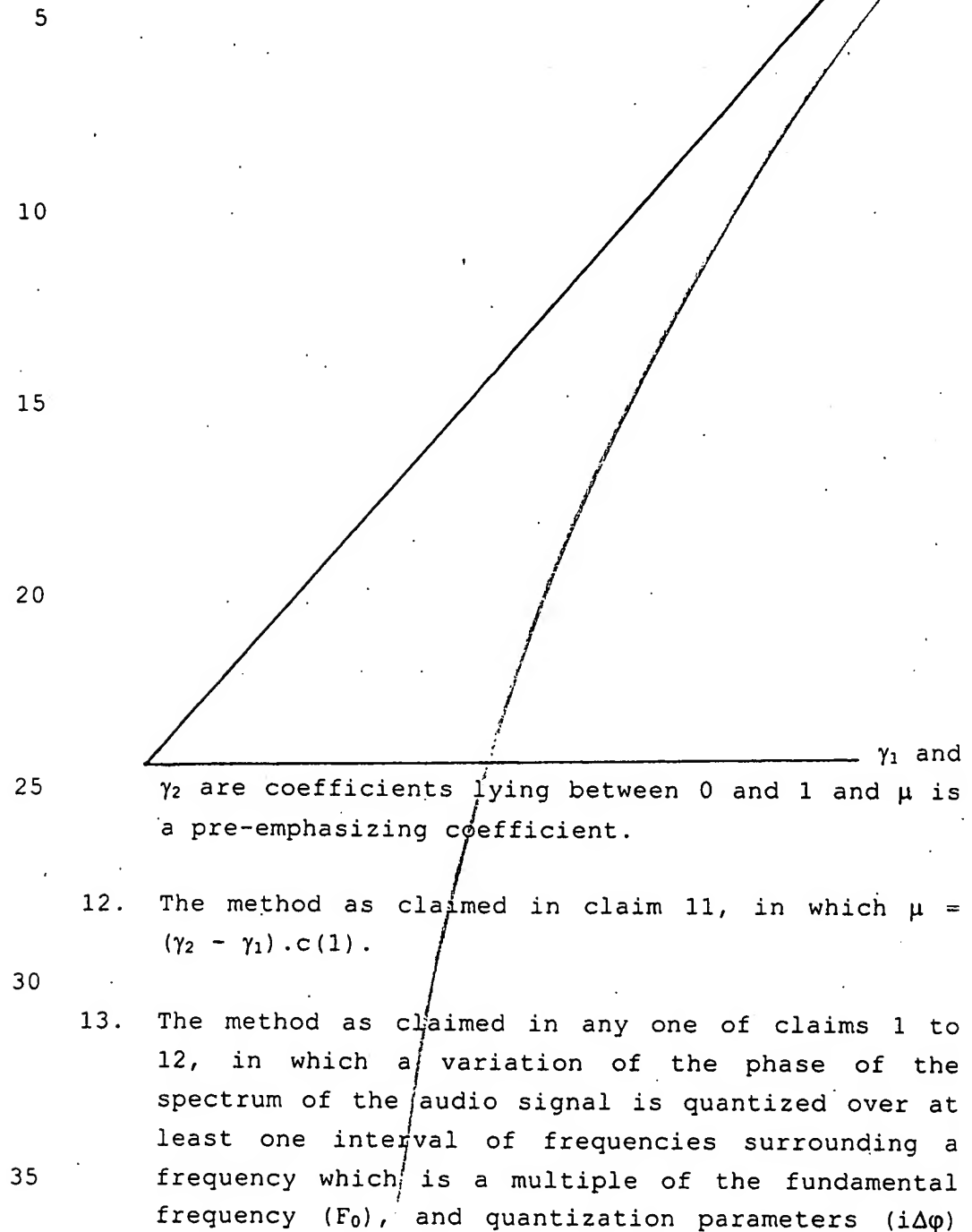
THIS PAGE BLANK (USPTO)

VII. Certain defects in the international application

The following defects in the form or contents of the international application have been noted:

Contrary to the requirement of PCT Rule 5.1(a)(ii), the relevant prior art disclosed in document D1 has not been indicated in the description, nor has this document been cited.

THIS PAGE BLANK (USPTO:



AMENDED SHEET

for said phase variation are included in the data for coding the harmonic component.

- 5 14. The method as claimed in claim 13, in which the quantized phase variation is represented by the slope of the phase of the spectrum at the frequency which is a multiple of the fundamental frequency (F_0).
- 10 15. The method as claimed in claim 13, in which the quantized phase variation represents a phase model stored in a dictionary determined by learning.
- 15 16. An audio coder, comprising programmed means for executing a method according to any one of claims 1 to 15.
- 20 17. A method of decoding a digital input stream (Φ) representing a coded audio signal, in which data for coding a fundamental frequency (F_0) of the audio signal and data for coding a harmonic component comprising quantization data for cepstral coefficients representing an upper spectral envelope (X_{sup}) of the audio signal are received in the digital input stream, and a spectral estimate (X_v) of the harmonic component is generated, its modulus being determined by said upper spectral envelope at the frequencies which are multiples of the fundamental frequency, characterized in that, to determine the phase (ϕ_v) of the spectral estimate of the harmonic component (X_v), the cepstral coefficients are transformed by liftering and smoothing in the cepstral domain, and minimum phases ($\phi(k)$) of the audio signal at
- 30 frequencies which are multiples
- 35

AMENDED SHEET

25. The method as claimed in claim 24, in which said threshold (d_{th}) is adapted as a function of the variations of the cepstral coefficients ($c_{xp}[n]$).
- 5 26. A method of decoding a digital input stream (Φ) representing a coded audio signal, in which data for coding a fundamental frequency (F_0) of the audio signal and data for coding a harmonic component comprising quantization data for cepstral coefficients representing an upper spectral envelope (X_{sup}) of the audio signal are received in the digital input stream, and a spectral estimate (X_v) of the harmonic component is generated, its modulus being determined by said upper spectral envelope at the frequencies which are multiples of the fundamental frequency and a phase, characterized in that a phase variation (ϕ_v) of the spectral estimate of the harmonic component is determined over at least one interval of frequencies surrounding a frequency which is a multiple of the fundamental frequency on the basis of quantization parameters ($i\Delta\phi$) included in the data for coding the harmonic component.
- 10 27. The method as claimed in claim 26, in which the phase variation determined over an interval of frequencies is a linear variation represented by a slope quantization parameter.
- 15 28. The method as claimed in claim 26, in which the phase variation determined over an interval of frequencies represents a phase model stored in a dictionary determined by learning.
- 20 25 30 35

29. An audio decoder, comprising programmed means for executing a method according to any one of claims 17 to 28.

AMENDED SHEET

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application No. :

U.S. National Serial No. :

Filed :

PCT International Application No. : PCT/FR00/01909

VERIFICATION OF A TRANSLATION

I, the below named translator, hereby declare that:

My name and post office address are as stated below;

That I am knowledgeable in the French language in which the below identified international application was filed, and that, to the best of my knowledge and belief, the English translation of the amended sheets of the international application No. PCT/FR00/01909 is a true and complete translation of the amended sheets of the above identified international application as filed.

I hereby declare that all the statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the patent application issued thereon.

Date: November 20, 2001



Full name of the translator :

David LAWSON

For and on behalf of RWS Group plc

Post Office Address :

Europa House, Marsham Way,
Gerrards Cross, Buckinghamshire,
England.

Express Mail No. EL857974156US

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application No. :

U.S. National Serial No. :

Filed :

PCT International Application No. : PCT/FR00/01909

VERIFICATION OF A TRANSLATION

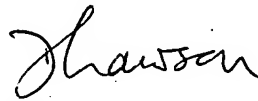
I, the below named translator, hereby declare that:

My name and post office address are as stated below;

That I am knowledgeable in the French language in which the below identified international application was filed, and that, to the best of my knowledge and belief, the English translation of the international application No. PCT/FR00/01909 is a true and complete translation of the above identified international application as filed.

I hereby declare that all the statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the patent application issued thereon.

Date: November 20, 2001



Full name of the translator :

David LAWSON

For and on behalf of RWS Group plc

Post Office Address :

Europa House, Marsham Way,
Gerrards Cross, Buckinghamshire,
England.

TRAITE DE COOPERATION EN MATIERE DE BREVETS

PCT

RAPPORT D'EXAMEN PRELIMINAIRE INTERNATIONAL

(article 36 et règle 70 du PCT)



Référence du dossier du déposant ou du mandataire BCT000055 BLO/EB/SDV	POUR SUITE A DONNER voir la notification de transmission du rapport d'examen préliminaire international (formulaire PCT/IPEA/416)	
Demande internationale n° PCT/FR00/01909	Date du dépôt international (<i>jour/mois/année</i>) 04/07/2000	Date de priorité (<i>jour/mois/année</i>) 05/07/1999
Classification internationale des brevets (CIB) ou à la fois classification nationale et CIB G10L19/02		
Déposant MATRA NORTEL COMMUNICATIONS et al.		

1. Le présent rapport d'examen préliminaire international, établi par l'administration chargée de l'examen préliminaire international, est transmis au déposant conformément à l'article 36.
2. Ce RAPPORT comprend 6 feuilles, y compris la présente feuille de couverture.
 - ☒ Il est accompagné d'ANNEXES, c'est-à-dire de feuilles de la description, des revendications ou des dessins qui ont été modifiées et qui servent de base au présent rapport ou de feuilles contenant des rectifications faites auprès de l'administration chargée de l'examen préliminaire international (voir la règle 70.16 et l'instruction 607 des Instructions administratives du PCT).

Ces annexes comprennent 2 feuilles.

3. Le présent rapport contient des indications relatives aux points suivants:

- I ☒ Base du rapport
- II ☐ Priorité
- III ☐ Absence de formulation d'opinion quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle
- IV ☐ Absence d'unité de l'invention
- V ☒ Déclaration motivée selon l'article 35(2) quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle; citations et explications à l'appui de cette déclaration
- VI ☐ Certains documents cités
- VII ☒ Irrégularités dans la demande internationale
- VIII ☐ Observations relatives à la demande internationale

Date de présentation de la demande d'examen préliminaire internationale 22/01/2001	Date d'achèvement du présent rapport 25.09.2001
Nom et adresse postale de l'administration chargée de l'examen préliminaire international:  Office européen des brevets D-80298 Munich Tél. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Fonctionnaire autorisé Bourdier, R N° de téléphone +49 89 2399 2130 

THIS PAGE BLANK (USPTO)

RAPPORT D'EXAMEN PRÉLIMINAIRE INTERNATIONAL

Demande internationale n° PCT/FR00/01909

I. Base du rapport

1. En ce qui concerne les **éléments** de la demande internationale (*les feuilles de remplacement qui ont été remises à l'office récepteur en réponse à une invitation faite conformément à l'article 14 sont considérées dans le présent rapport comme "initialement déposées" et ne sont pas jointes en annexe au rapport puisqu'elles ne contiennent pas de modifications (règles 70.16 et 70.17)*):

Description, pages:

1-32 version initiale

Revendications, N°:

1-10,11 (partie), version initiale
17 (partie),
18-24 -----

11 (partie),12-16, reçue(s) le 07/06/2001 avec la lettre du 06/06/2001
17 (partie),
25-29

Dessins, feuilles:

1/13-13/13 version initiale

2. En ce qui concerne la **langue**, tous les éléments indiqués ci-dessus étaient à la disposition de l'administration ou lui ont été remis dans la langue dans laquelle la demande internationale a été déposée, sauf indication contraire donnée sous ce point.

Ces éléments étaient à la disposition de l'administration ou lui ont été remis dans la langue suivante: , qui est :

- ☐ la langue d'une traduction remise aux fins de la recherche internationale (selon la règle 23.1(b)).
- ☐ la langue de publication de la demande internationale (selon la règle 48.3(b)).
- ☐ la langue de la traduction remise aux fins de l'examen préliminaire internationale (selon la règle 55.2 ou 55.3).

3. En ce qui concerne les **séquences de nucléotides ou d'acide aminés** divulguées dans la demande internationale (le cas échéant), l'examen préliminaire internationale a été effectué sur la base du listage des séquences :

- ☐ contenu dans la demande internationale, sous forme écrite.
- ☐ déposé avec la demande internationale, sous forme déchiffrable par ordinateur.
- ☐ remis ultérieurement à l'administration, sous forme écrite.
- ☐ remis ultérieurement à l'administration, sous forme déchiffrable par ordinateur.

THIS PAGE BLANK (USPTO)

**RAPPORT D'EXAMEN
PRÉLIMINAIRE INTERNATIONAL**

Demande internationale n° PCT/FR00/01909

- ☐ La déclaration, selon laquelle le listage des séquences par écrit et fourni ultérieurement ne va pas au-delà de la divulgation faite dans la demande telle que déposée, a été fournie.
- ☐ La déclaration, selon laquelle les informations enregistrées sous déchiffrable par ordinateur sont identiques à celles du listage des séquences Présenté par écrit, a été fournie.

4. Les modifications ont entraîné l'annulation :

- ☐ de la description, pages :
- ☐ des revendications, n°s :
- ☐ des dessins, feuilles :

5. ☐ Le présent rapport a été formulé abstraction faite (de certaines) des modifications, qui ont été considérées comme allant au-delà de l'exposé de l'invention tel qu'il a été déposé, comme il est indiqué ci-après (règle 70.2(c)) :

(Toute feuille de remplacement comportant des modifications de cette nature doit être indiquée au point 1 et annexée au présent rapport)

6. Observations complémentaires, le cas échéant :

V. Déclaration motivée selon l'article 35(2) quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle; citations et explications à l'appui de cette déclaration

1. Déclaration

Nouveauté	Oui : Revendications 1-29
	Non : Revendications
Activité inventive	Oui : Revendications 1-29
	Non : Revendications
Possibilité d'application industrielle	Oui : Revendications 1-29
	Non : Revendications

2. Citations et explications
voir feuille séparée

VII. Irrégularités dans la demande internationale

Les irrégularités suivantes, concernant la forme ou le contenu de la demande internationale, ont été constatées :
voir feuille séparée

THIS PAGE BLANK (USPTO)

CONCERNANT LA SECTION V

1. L'état de la technique le plus pertinent pour la demande est considéré comme étant D1: AHMADI S ET AL: 'New techniques for sinusoidal coding of speech at 2400 bps' CONFERENCE RECORD OF THIRTIETH ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS (CAT. NO.96CB36004), PACIFIC GROVE, CA, USA, 3-6 NOV. 1996, pages 770-774 vol.1, XP002138769 1997, Los Alamitos, CA, USA, IEEE Comput. Soc. Press, USA ISBN: 0-8186-7646-9.
2. En utilisant le libellé de la revendication 1, D1 divulgue un procédé de codage d'un signal audio (page 770, 3 premières lignes de l'Introduction) dans lequel on estime une fréquence fondamentale (page 770, colonne de gauche, 2 dernières lignes), on détermine un spectre par transformation dans le domaine fréquentiel (LPC, page 770, colonne de droite, lignes 16-25), et on inclut dans un flot de données numériques des données représentatives d'amplitude spectrales aux multiples de la fréquence fondamentale (A^{km} , Φ^{km} , page 770, colonne de droite, lignes 20-23, voir aussi chapitre 3, premier paragraphe, page 771). Dans D1, les données représentatives A^{km} sont obtenues en échantillonnant l'enveloppe LPC aux multiples de la fréquence fondamentale dont la valeur a été déterminée par le biais de coefficients cepstraux calculés en transformant dans le domaine cepstral le spectre comprimé (par application d'un logarithme, comme il est usuel) du signal audio. D'autre part, les données contiennent une information de phase Φ^{km} qui représente la phase du spectre aux fréquences multiples de la fréquence fondamentale (voir page 770, colonne de droite, lignes 4-15, Φ^{km} est associé à A^{km} , amplitude de la raie multiple m du segment k).

Par conséquent, seule la caractéristique selon laquelle c'est une enveloppe supérieure du spectre audio qui est transformée pour obtenir les coefficients cepstraux, est nouvelle par rapport à D1.

3. Dans D1, le cepstre est calculé afin de déterminer l'information de fréquence fondamentale, aisément lisible sur ledit cepstre, comme il est bien connu. Il est nécessaire pour cela de calculer la transformation inverse du logarithme du module du spectre, brut. Si en effet, on ne considérait que l'enveloppe du spectre, l'information des raies spectrales disparaîtrait et le calcul

THIS PAGE BLANK (USPTO)

de fréquence fondamentale ne serait plus possible sur le cepstre résultant. Par conséquent, il n'y aurait aucune raison pour l'homme du métier de modifier l'enseignement de D1 pour y introduire le calcul cepstral à partir d'une enveloppe supérieure du spectre audio, comme revendiqué.

Cette caractéristique nouvelle, qui dans la demande sert à d'autres fins, doit donc être considérée comme inventive.

4. Les revendications 2 à 15 sont considérées comme étant nouvelles et inventives pour les mêmes raisons que la revendication 1 dont elles dépendent.
5. Le codeur audio de la revendication 16, programmé pour exécuter un procédé jugé nouveau et inventif, est également considéré comme étant nouveau et inventif.
6. Les caractéristiques de détermination de la phase comme faisant partie du décodage, telles que revendiquées à la revendication 17, ne sont ni connues, ni rendues évidentes par D1 ou par un autre document. La revendication 17 est donc considérée comme étant nouvelle et inventive.
7. Les revendications 18 à 25 sont considérées comme étant nouvelles et inventives pour la même raison que la revendication 17, dont elles dépendent.
8. Les caractéristiques de détermination de variation de la phase comme faisant partie du décodage, telles que revendiquées à la revendication 26, ne sont ni connues, ni rendues évidentes par D1 ou par un autre document. La revendication 26 est donc considérée comme étant nouvelle et inventive.
9. Le décodeur audio de la revendication 29, programmé pour exécuter un procédé jugé nouveau et inventif, est également considéré comme étant nouveau et inventif.
10. Toutes les revendications sont susceptibles d'application industrielle

THIS PAGE BLANK (USPT[®])

CONCERNANT LA SECTION VII

1. Contrairement à ce qu'exige la règle 5.1 a) ii) PCT, la description n'indique pas l'état de la technique antérieure pertinent exposé dans le document D1 et ne cite pas ce document.

THIS PAGE BLANK (USPTO)

- 35 -

γ_1 et γ_2 sont des coefficients compris entre 0 et 1 et μ est un coefficient de préaccentuation.

12. Procédé selon la revendication 11, dans lequel $\mu = (\gamma_2 - \gamma_1) \cdot c(1)$.

13. Procédé selon l'une quelconque des revendications 1 à 12, dans lequel on quantifie une variation de la phase du spectre du signal audio sur au moins un intervalle de fréquences entourant une fréquence multiple de la fréquence fondamentale (F_0), et on inclut des paramètres ($i\Delta\phi$) de quantification de ladite variation de phase dans les données de codage de la composante harmonique.

14. Procédé selon la revendication 13, dans lequel la variation de phase quantifiée est représentée par la pente de la phase du spectre à la fréquence multiple de la fréquence fondamentale (F_0).

15. Procédé selon la revendication 13, dans lequel la variation de phase quantifiée représente un modèle de phase mémorisé dans un dictionnaire déterminé par apprentissage.

16. Codeur audio, comprenant des moyens programmés pour exécuter un procédé selon l'une quelconque des revendications 1 à 15.

17. Procédé de décodage d'un flux numérique d'entrée (Φ) représentant un signal audio codé, dans lequel on reçoit dans le flux numérique d'entrée des données de codage d'une fréquence fondamentale (F_0) du signal audio et des données de codage d'une composante harmonique comprenant des données de quantification de coefficients cepstraux représentant une enveloppe spectrale supérieure (X_{sup}) du signal audio, et on génère une estimation spectrale (X_v) de la composante harmonique dont le module est déterminé par ladite enveloppe spectrale supérieure aux fréquences multiples de la fréquence fondamentale, caractérisé en ce que, pour déterminer la phase (ϕ_v) de l'estimation spectrale de la composante harmonique (X_v), on transforme les coefficients cepstraux par liftrage et lissage dans le domaine cepstral, et on calcule des phases minimales ($\phi(k)$) du signal audio à des fréquences multiples

FEUILLE MODIFIEE

THIS PAGE BLANK (USPTO)

- 37 -

25. Procédé selon la revendication 24, dans lequel ledit seuil (d_{th}) est adapté en fonction des variations des coefficients cepstraux ($c_{xp}[n]$).

26. Procédé de décodage d'un flux numérique d'entrée (Φ) représentant un signal audio codé, dans lequel on reçoit dans le flux numérique d'entrée des données de codage d'une fréquence fondamentale (F_0) du signal audio et des données de codage d'une composante harmonique comprenant des données de quantification de coefficients cepstraux représentant une enveloppe spectrale supérieure (X_{sup}) du signal audio, et on génère une estimation spectrale (X_v) de la composante harmonique ayant un module déterminé par ladite enveloppe spectrale supérieure aux fréquences multiples de la fréquence fondamentale et une phase, caractérisé en ce qu'on détermine une variation de la phase (ϕ_v) de l'estimation spectrale de la composante harmonique sur au moins un intervalle de fréquences entourant une fréquence multiple de la fréquence fondamentale sur la base de paramètres de quantification ($i\Delta\phi$) inclus dans les données de codage de la composante harmonique.

27. Procédé selon la revendication 26, dans lequel la variation de phase déterminée sur un intervalle de fréquences est une variation linéaire représentée par un paramètre de quantification de pente.

28. Procédé selon la revendication 26, dans lequel la variation de phase déterminée sur un intervalle de fréquences représente un modèle de phase mémorisé dans un dictionnaire déterminé par apprentissage.

29. Décodeur audio, comprenant des moyens programmés pour exécuter un procédé selon l'une quelconque des revendications 17 à 28.

FEUILLE MODIFIÉE

THIS PAGE BLANK (USPTO)

PCT

RAPPORT DE RECHERCHE INTERNATIONALE

(article 18 et règles 43 et 44 du PCT)

Référence du dossier du déposant ou du mandataire BCT000055	POUR SUITE A DONNER voir la notification de transmission du rapport de recherche internationale (formulaire PCT/ISA/220) et, le cas échéant, le point 5 ci-après	
Demande internationale n° PCT/FR 00/ 01909	Date du dépôt international(jour/mois/année) 04/07/2000	(Date de priorité (la plus ancienne) (jour/mois/année) 05/07/1999
Déposant MATRA NORTEL COMMUNICATIONS		

Le présent rapport de recherche internationale, établi par l'administration chargée de la recherche internationale, est transmis au déposant conformément à l'article 18. Une copie en est transmise au Bureau international.

Ce rapport de recherche internationale comprend 3 feuilles.



Il est aussi accompagné d'une copie de chaque document relatif à l'état de la technique qui y est cité.

1. Base du rapport

- a. En ce qui concerne la **langue**, la recherche internationale a été effectuée sur la base de la demande internationale dans la langue dans laquelle elle a été déposée, sauf indication contraire donnée sous le même point.



la recherche internationale a été effectuée sur la base d'une traduction de la demande internationale remise à l'administration.

- b. En ce qui concerne les **séquences de nucléotides ou d'acides aminés** divulguées dans la demande internationale (le cas échéant), la recherche internationale a été effectuée sur la base du listage des séquences :



contenu dans la demande internationale, sous forme écrite.



déposée avec la demande internationale, sous forme déchiffrable par ordinateur.



remis ultérieurement à l'administration, sous forme écrite.



remis ultérieurement à l'administration, sous forme déchiffrable par ordinateur.



La déclaration, selon laquelle le listage des séquences présenté par écrit et fourni ultérieurement ne va pas au-delà de la divulgation faite dans la demande telle que déposée, a été fournie.



La déclaration, selon laquelle les informations enregistrées sous forme déchiffrable par ordinateur sont identiques à celles du listage des séquences présenté par écrit, a été fournie.

2. ☐ Il a été estimé que certaines revendications ne pouvaient pas faire l'objet d'une recherche (voir le cadre I).

3. ☐ Il y a absence d'unité de l'invention (voir le cadre II).

4. En ce qui concerne le titre,



le texte est approuvé tel qu'il a été remis par le déposant.



Le texte a été établi par l'administration et a la teneur suivante:

CODAGE ET DECODAGE AUDIO AVEC COMPOSANTS HARMONIQUES ET PHASE MINIMALE

5. En ce qui concerne l'abrégé,



le texte est approuvé tel qu'il a été remis par le déposant



le texte (reproduit dans le cadre III) a été établi par l'administration conformément à la règle 38.2b). Le déposant peut présenter des observations à l'administration dans un délai d'un mois à compter de la date d'expédition du présent rapport de recherche internationale.

6. La figure des dessins à publier avec l'abrégé est la Figure n°



suggérée par le déposant.



parce que le déposant n'a pas suggéré de figure.



parce que cette figure caractérise mieux l'invention.

25



Aucune des figures n'est à publier.

THIS PAGE BLANK (USPTO)

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/FR 00/01909

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G10L19/02

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
INSPEC, EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	AHMADI S ET AL: "New techniques for sinusoidal coding of speech at 2400 bps" CONFERENCE RECORD OF THIRTIETH ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS (CAT. NO.96CB36004), PACIFIC GROVE, CA, USA, 3-6 NOV. 1996, pages 770-774 vol.1, XP002138769 1997, Los Alamitos, CA, USA, IEEE Comput. Soc. Press, USA ISBN: 0-8186-7646-9	1,2,16, 17,29
Y	abstract page 770, right-hand column, line 16-38 * pages 772,773 : section "4 Sinusoidal Phase Coding" * ----- -/-	13,26

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

7 September 2000

Date of mailing of the international search report

14/09/2000

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Quélavoine, R

INTERNATIONAL SEARCH REPORT

In International Application No

PCT/FR 00/01909

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 0 709 827 A (MITSUBISHI ELECTRIC CORP) 1 May 1996 (1996-05-01) abstract column 2, line 54 -column 3, line 53 -----	13,26
A	US 5 651 092 A (ISHII ET AL) 22 July 1997 (1997-07-22) claim 1 column 2, line 21-29 column 2, line 39-47 -----	1,4,13, 17,26
A	US 5 701 391 A (PAN ET AL) 23 December 1997 (1997-12-23) abstract column 1, line 62 -column 2, line 15 column 2, line 26-39 column 8, line 15-31 -----	1,2,17

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. l. Application No

PCT/FR 00/01909

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0709827 A	01-05-1996	JP 8123494 A	17-05-1996
		CA 2160749 A	29-04-1996
		CN 1126869 A	17-07-1996
		KR 169020 B	20-03-1999
		US 5724480 A	03-03-1998
US 5651092 A	22-07-1997	JP 6332496 A	02-12-1994
		CA 2122853 A,C	22-11-1994
		CA 2214585 A	22-11-1994
		DE 69420183 D	30-09-1999
		DE 69420183 T	09-12-1999
		EP 0626674 A	30-11-1994
		EP 0854469 A	22-07-1998
		US 5596675 A	21-01-1997
US 5701391 A	23-12-1997	AU 7469296 A	22-05-1997
		WO 9716820 A	09-05-1997

THIS PAGE BLANK (USPTO)

RAPPORT DE RECHERCHE INTERNATIONALE

De de Internationale No

PCT/FR 00/01909

A. CLASSEMENT DE L'OBJET DE LA DEMANDE
CIB 7 G10L19/02

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)
CIB 7 G10L

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)
INSPEC, EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	AHMADI S ET AL: "New techniques for sinusoidal coding of speech at 2400 bps" CONFERENCE RECORD OF THIRTIETH ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS (CAT. NO.96CB36004), PACIFIC GROVE, CA, USA, 3-6 NOV. 1996, pages 770-774 vol.1, XP002138769 1997, Los Alamitos, CA, USA, IEEE Comput. Soc. Press, USA ISBN: 0-8186-7646-9	1,2,16, 17,29
Y	abrégé page 770, colonne de droite, ligne 16-38 * pages 772,773 : section "4 Sinusoidal Phase Coding" *	13,26

☒ Voir la suite du cadre C pour la fin de la liste des documents

☒ Les documents de familles de brevets sont indiqués en annexe

* Catégories spéciales de documents cités:

- *A* document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- *E* document antérieur, mais publié à la date de dépôt international ou après cette date
- *L* document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- *O* document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- *P* document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

T document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

- *X* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément
- *Y* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier
- *&* document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

7 septembre 2000

Date d'expédition du présent rapport de recherche internationale

14/09/2000

Nom et adresse postale de l'administration chargée de la recherche internationale
Office Européen des Brevets, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Quélavoine, R

RAPPORT DE RECHERCHE INTERNATIONALE

De de Internationale No
PCT/FR 00/01909

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
Y	EP 0 709 827 A (MITSUBISHI ELECTRIC CORP) 1 mai 1996 (1996-05-01) abrégé colonne 2, ligne 54 -colonne 3, ligne 53 _____	13,26
A	US 5 651 092 A (ISHII ET AL) 22 juillet 1997 (1997-07-22) revendication 1 colonne 2, ligne 21-29 colonne 2, ligne 39-47 _____	1,4,13, 17,26
A	US 5 701 391 A (PAN ET AL) 23 décembre 1997 (1997-12-23) abrégé colonne 1, ligne 62 -colonne 2, ligne 15 colonne 2, ligne 26-39 colonne 8, ligne 15-31 _____	1,2,17

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

De: de Internationale No

PCT/FR 00/01909

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
EP 0709827 A	01-05-1996	JP 8123494 A	17-05-1996
		CA 2160749 A	29-04-1996
		CN 1126869 A	17-07-1996
		KR 169020 B	20-03-1999
		US 5724480 A	03-03-1998
US 5651092 A	22-07-1997	JP 6332496 A	02-12-1994
		CA 2122853 A,C	22-11-1994
		CA 2214585 A	22-11-1994
		DE 69420183 D	30-09-1999
		DE 69420183 T	09-12-1999
		EP 0626674 A	30-11-1994
		EP 0854469 A	22-07-1998
		US 5596675 A	21-01-1997
US 5701391 A'	23-12-1997	AU 7469296 A	22-05-1997
		WO 9716820 A	09-05-1997

THIS PAGE BLANK (USPTO)

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
11 janvier 2001 (11.01.2001)

PCT

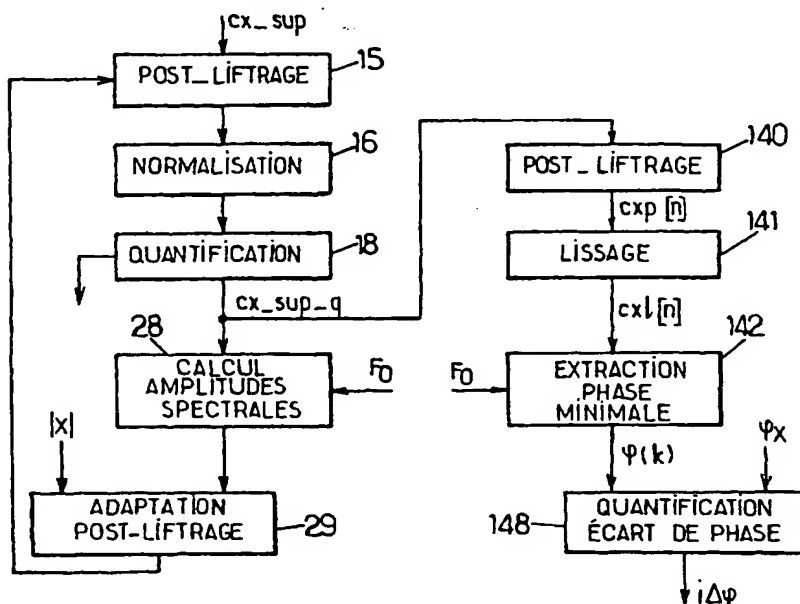
(10) Numéro de publication internationale
WO 01/03121 A1

- (51) Classification internationale des brevets⁷: G10L 19/02 (71) Déposant (pour tous les États désignés sauf US): MATRA NORTEL COMMUNICATIONS [FR/FR]; 50, rue du Président Sadate, F-29100 Quimper (FR).
- (21) Numéro de la demande internationale: PCT/FR00/01909 (72) Inventeurs; et (75) Inventeurs/Déposants (pour US seulement): CAPMAN, François [FR/FR]; 47, rue des Etats-Généraux, F-78000 Versailles (FR). MURGIA, Carlo [IT/US]; 458 Saint Vincent, Irvine, CA 92618 (US).
- (22) Date de dépôt international: 4 juillet 2000 (04.07.2000)
- (25) Langue de dépôt: français (74) Mandataires: LOISEL, Bertrand etc.; Cabinet Plasseraud, 84, rue d'Amsterdam, F-75440 Paris Cedex 09 (FR).
- (26) Langue de publication: français (81) États désignés (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE,
- (30) Données relatives à la priorité: 99/08633 5 juillet 1999 (05.07.1999) FR

[Suite sur la page suivante]

(54) Title: ENCODING AND DECODING WITH HARMONIC COMPONENTS AND MINIMUM PHASE

(54) Titre: CODAGE ET DECODAGE AUDIO AVEC COMPOSANTS HARMONIQUES ET PHASE MINIMALE



15,140... POST-LIFTERING
18... QUANTIZATION
28... SPECTRAL AMPLITUDE CALCULATION
29... POST-LIFTERING ADAPTATION
141... SMOOTHING
142... MINIMUM PHASE RETRIEVAL
148... PHASE DIFFERENCE QUANTIZATION

(57) Abstract: The invention concerns a method wherein the encoder estimates a fundamental frequency (F_0) of an audio signal, determines a spectrum of the audio signal by a transform in the frequency domain of a frame of the audio signal, and includes in the digital flow transmitted to the decoder data encoding a harmonic component of the audio signal, comprising data representing spectral amplitudes associated with multiple frequencies of the estimated fundamental frequency. The data encoding the harmonic component further comprise, for at least one of the multiple frequencies of the estimated fundamental frequency, data ($i\Delta\phi$) concerning the phase of the audio signal spectrum in the neighbourhood of said multiple frequency.

Express Mail No. EL857974156US

[Suite sur la page suivante]



WO 01/03121 A1



DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Publiée:

— Avec rapport de recherche internationale.

(84) États désignés (*régional*): brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), brevet eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU,

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

(57) Abrégé: Le codeur estime une fréquence fondamentale (F_0) du signal audio, détermine un spectre du signal audio par une transformée dans le domaine fréquentiel d'une trame du signal audio, et inclut dans le flux numérique transmis au décodeur des données de codage d'une composante harmonique du signal audio, comprenant des données représentatives d'amplitudes spectrales associées à des fréquences multiples de la fréquence fondamentale estimée. Les données de codage de la composante harmonique comprennent en outre, pour au moins une des fréquences multiples de la fréquence fondamentale estimée, des données ($i\Delta\phi$) relatives à la phase du spectre du signal audio au voisinage de cette fréquence multiple.

CODAGE ET DECODAGE AUDIO AVEC COMPOSANTS HARMONIQUES ET PHASE MINIMALE

La présente invention concerne le domaine du codage des signaux audio. Elle s'applique en particulier, mais non exclusivement, au codage de la parole, en bande étroite ou en bande élargie, dans diverses gammes de débit de codage.

La conception d'un codec audio vise principalement à fournir un bon compromis entre le débit du flux transmis par le codeur et la qualité du signal audio que le décodeur est capable de reconstruire à partir de ce flux.

Dans cette optique, il a notamment été développé des familles de codeurs reposant sur une analyse du signal audio dans le domaine spectral : le codeur estime une fréquence fondamentale du signal, représentant sa hauteur tonale (« pitch »), et l'analyse spectrale consiste à déterminer des paramètres représentant la structure harmonique du signal aux fréquences qui sont des multiples entiers de cette fréquence fondamentale. Une modélisation de la composante non-harmonique, ou non-voisée, peut également être effectuée dans le domaine spectral. Les paramètres transmis au décodeur représentent typiquement le module du spectre des composantes voisées et non-voisées. Il s'y ajoute des informations représentant soit des décisions voisé/non-voisé relatives à différentes portions de spectre, soit des informations sur la probabilité de voisement du signal, permettant au décodeur de déterminer dans quelles portions du spectre il doit utiliser la composante voisée ou la composante non-voisée.

Ces familles de codeurs comprennent les codeurs du type MBE (« Multi-Band Excitation »), ou encore les codeurs de type STC (« Sinusoidal Transform Coder »). A titre de référence, on peut citer les brevets US 4 856 068, 4 885 790, 4 937 873, 5 054 072, 5 081 681, 5 195 166, 5 216 747, 5 226 084, 5 226 108, 5 247 579, 5 473 727, 5 517 511, 5 630 011, 5 630 012, 5 649 050, 5 651 093, 5 664 051, 5 664 052, 5 684 926, 5 701 390, 5 715 365, 5 749 065, 5 752 222, 5 765 127, 5 774 837 et 5 890 108.

Un but de la présente invention est de permettre, dans un schéma de codage à analyse dans le domaine spectral, d'améliorer la modélisation des phases du spectre du signal par le décodeur.

L'invention propose ainsi un procédé de codage d'un signal audio, dans lequel on estime une fréquence fondamentale du signal audio, on détermine un spectre du signal audio par une transformée dans le domaine

fréquentiel d'une trame du signal audio, et on inclut dans un flux numérique de sortie des données de codage d'une composante harmonique du signal audio, comprenant des données représentatives d'amplitudes spectrales associées à des fréquences multiples de la fréquence fondamentale estimée. Selon
5 l'invention, lesdites données représentatives d'amplitudes spectrales associées à des fréquences multiples de la fréquence fondamentale sont obtenues au moyen de coefficients cepstraux calculés en transformant dans le domaine cepstral une enveloppe supérieure comprimée du spectre du signal audio, et les données de codage de la composante harmonique comprennent en outre,
10 pour au moins une des fréquences multiples de la fréquence fondamentale estimée, des données relatives à la phase du spectre du signal audio au voisinage de ladite fréquence multiple.

Un autre aspect de la présente invention se rapporte à un procédé de décodage d'un flux numérique d'entrée représentant un signal audio codé,
15 dans lequel on reçoit dans le flux numérique d'entrée des données de codage d'une fréquence fondamentale du signal audio et des données de codage d'une composante harmonique comprenant des données de quantification de coefficients cepstraux représentant une enveloppe spectrale supérieure du signal audio, et on génère une estimation spectrale de la composante
20 harmonique dont le module est déterminé par ladite enveloppe spectrale supérieure aux fréquences multiples de la fréquence fondamentale. Selon l'invention, pour déterminer la phase de l'estimation spectrale de la composante harmonique, on transforme les coefficients cepstraux par liftrage et lissage dans le domaine cepstral, et on calcule des phases minimales du
25 signal audio à des fréquences multiples de la fréquence fondamentale sur la base des coefficients cepstraux transformés. Selon l'invention, on peut également déterminer une variation de la phase de l'estimation spectrale de la composante harmonique sur au moins un intervalle de fréquences entourant une fréquence multiple de la fréquence fondamentale sur la base de
30 paramètres de quantification inclus dans les données de codage de la composante harmonique.

L'invention propose également un codeur et un décodeur audio comprenant des moyens pour la mise en œuvre des procédés ci-dessus.

D'autres particularités et avantages de la présente invention
35 apparaîtront dans la description ci-après d'exemples de réalisation non limitatifs, en référence aux dessins annexés, dans lesquels :

- la figure 1 est un schéma synoptique d'un codeur audio selon l'invention ;
- les figures 2 et 3 sont des diagrammes illustrant la formation des trames de signal audio dans le codeur de la figure 1 ;
- les figures 4 et 5 sont des graphiques montrant un exemple de spectre du signal audio et illustrant l'extraction des enveloppes supérieure et inférieure de ce spectre ;
- la figure 6 est un schéma synoptique d'un exemple de moyens de quantification utilisables dans le codeur de la figure 1 ;
- la figure 7 est un schéma synoptique de moyens utilisables pour extraire des paramètres se rapportant à la phase de la composante non-harmonique dans une variante du codeur de la figure 1 ;
- la figure 8 est un schéma synoptique d'un décodeur audio correspondant au codeur de la figure 1 ;
- la figure 9 est un organigramme d'un exemple de procédure de lissage de coefficients spectraux et d'extraction de phases minimales mise en œuvre dans le décodeur de la figure 8 ;
- la figure 10 est un schéma synoptique de modules d'analyse et de mixage spectral de composantes harmonique et non-harmonique du signal audio ;
- les figures 11 à 13 sont des graphiques montrant des exemples de fonctions non-linéaires utilisables dans le module d'analyse de la figure 10 ;
- les figures 14 et 15 sont des diagrammes illustrant une façon de procéder à la synthèse temporelle des trames de signal dans le décodeur de la figure 8 ;
- les figures 16 et 17 sont des graphiques montrant des fonctions de fenêtrage utilisables dans la synthèse des trames suivant les figures 14 et 15 ;
- les figures 18 et 19 sont des schémas synoptiques de moyens d'interpolation utilisables dans une variante de réalisation du codeur et du décodeur ;
- la figure 20 est un schéma synoptique de moyens d'interpolation utilisables dans une autre variante de réalisation du codeur ;
- les figures 21 et 22 sont des diagrammes illustrant une autre façon de procéder à la synthèse temporelle des trames de signal dans le décodeur de la figure 8, à l'aide d'une interpolation de paramètres ;

- les figures 23 à 25 sont des schémas synoptiques de variantes de moyens de post-traitement des coefficients cepstraux représentant l'enveloppe supérieure du spectre du signal dans le codeur de la figure 1 ; et
- 5 - la figure 26 est un schéma synoptique partiel d'un décodeur associé à un codeur selon la figure 25.

Le codeur et le décodeur décrits ci-après sont des circuits numériques qui peuvent, comme il est usuel dans le domaine du traitement des signaux audio, être réalisés par programmation d'un processeur de signal numérique (DSP) ou d'un circuit intégré d'application spécifique (ASIC).

10 Le codeur audio représenté sur la figure 1 traite un signal audio d'entrée x qui, dans l'exemple non-limitatif considéré ci-après, est un signal de parole. Le signal x est disponible sous forme numérique, par exemple à une fréquence d'échantillonnage F_e de 8 kHz. Il est par exemple délivré par un

15 convertisseur analogique-numérique traitant le signal de sortie amplifié d'un microphone. Le signal d'entrée x peut également être formé à partir d'une autre version, analogique ou numérique, codée ou non, du signal de parole.

Le codeur comprend un module 1 qui forme des trames successives de signal audio pour les différents traitements effectués, et un multiplexeur de

20 sortie 6 qui délivre un flux de sortie Φ contenant pour chaque trame des jeux de paramètres de quantification à partir desquels un décodeur sera capable de synthétiser une version décodée du signal audio.

La structure des trames est illustrée par les figures 2 et 3. Chaque trame 2 est composée d'un nombre N d'échantillons consécutifs du signal audio x . Les trames successives présentent des décalages temporels mutuels

25 correspondant à M échantillons, de sorte que leur recouvrement est de $L = N - M$ échantillons du signal. Dans l'exemple considéré, où $N = 256$, $M = 160$ et $L = 96$, la durée des trames 2 est de $N/F_e = 32$ ms, et une trame est formée toutes les $M/F_e = 20$ ms.

30 De façon classique, le module 1 multiplie les échantillons de chaque trame 2 par une fonction de fenêtrage f_A , de préférence choisie pour ses bonnes propriétés spectrales. Les échantillons $x(i)$ de la trame étant numérotés de $i = 0$ à $i = N - 1$, la fenêtre d'analyse $f_A(i)$ peut ainsi être une fenêtre de Hamming, d'expression :

$$f_A(i) = 0,54 + 0,46 \cdot \cos\left(2\pi \frac{i - (N-1)/2}{N}\right) \quad (1)$$

ou une fenêtre de Hanning d'expression :

$$f_A(i) = \frac{1}{2} \left(1 + \cos\left(2\pi \frac{i - (N-1)/2}{N}\right) \right) \quad (2)$$

ou encore une fenêtre de Kaiser, d'expression :

$$f_A(i) = \frac{I_0\left(\alpha \sqrt{1 - \left(\frac{i - (N-1)/2}{N}\right)^2}\right)}{I_0(\alpha)} \quad (3)$$

où α est un coefficient par exemple égal à 6, et $I_0(\cdot)$ désigne la fonction de Bessel d'indice 0.

Le codeur de la figure 1 procède à une analyse du signal audio dans le domaine spectral. Il comporte un module 3 qui calcule la transformée de Fourier rapide (TFR) de chaque trame de signal. La trame de signal est mise en forme avant d'être soumise au module de TFR 3 : le module 1 lui adjoint $N = 256$ échantillons à zéro afin d'obtenir la résolution maximale de la transformée de Fourier, et il effectue d'autre part une permutation circulaire des $2N = 512$ échantillons afin de compenser les effets de phase résultant de la fenêtre d'analyse. Cette modification de la trame est illustrée par la figure 3. La trame dont on calcule la transformée de Fourier rapide sur $2N = 512$ points commence par les $N/2 = 128$ derniers échantillons pondérés de la trame, suivis par les $N = 256$ échantillons à zéro, et se termine par les $N/2 = 128$ premiers échantillons pondérés de la trame.

Le module de TFR 3 obtient le spectre du signal pour chaque trame, dont le module et la phase sont respectivement notés $|X|$ et φ_X , ou $|X(i)|$ et $\varphi_X(i)$ pour les index de fréquence $i = 0$ à $i = 2N-1$ (grâce à la symétrie de la transformée de Fourier et des trames, on peut se limiter aux valeurs pour $0 \leq i < N$).

Un détecteur de fréquence fondamentale 4 estime pour chaque trame de signal une valeur de la fréquence fondamentale F_0 . Le détecteur 4 peut appliquer toute méthode connue d'analyse du signal de parole de la trame pour estimer la fréquence fondamentale F_0 , par exemple une méthode basée sur la fonction d'autocorrélation ou la fonction AMDF, éventuellement précédée d'un

module de blanchiment par prédiction linéaire. L'estimation peut également être effectuée dans le domaine spectral ou dans le domaine cepstral. Une autre possibilité est d'évaluer les intervalles de temps entre les ruptures consécutives du signal de parole attribuables à des fermetures de la glotte du locuteur intervenant pendant la durée de la trame. Des méthodes bien connues utilisables pour détecter de telles micro-ruptures sont décrites dans les articles suivants : M. Basseville et al., « Sequential detection of abrupt changes in spectral characteristics of digital signals » (IEEE Trans. on Information Theory, 1983, Vol. IT-29, n° 5, pages 708-723) ; R. Andre-Obrecht, « A new statistical approach for the automatic segmentation of continuous speech signals » (IEEE Trans. on Acous., Speech and Sig. Proc., Vol. 36, N°1, janvier 1988) ; et C. MURGIA et al., « An algorithm for the estimation of glottal closure instants using the sequential detection of abrupt changes in speech signals » (Signal Processing VII, 1994, pages 1685-1688).

La fréquence fondamentale estimée F_0 fait l'objet d'une quantification, par exemple scalaire, par un module 5, qui fournit au multiplexeur de sortie 6 un index iF de quantification de la fréquence fondamentale pour chaque trame du signal.

Le codeur utilise des modélisations paramétriques cepstrales pour représenter une enveloppe supérieure et une enveloppe inférieure du spectre du signal audio. La première étape de la transformation cepstrale consiste à appliquer au module du spectre du signal une fonction de compression spectrale, qui peut être une fonction logarithmique ou en racine. Le module 8 du codeur opère ainsi, pour chaque valeur $X(i)$ du spectre du signal ($0 \leq i < N$), la transformation suivante :

$$LX(i) = \text{Log}(|X(i)|) \quad (4)$$

dans le cas d'une compression logarithmique ou

$$LX(i) = |X(i)|^\gamma \quad (5)$$

dans le cas d'une compression en racine, γ étant un exposant compris entre 0 et 1.

Le spectre comprimé LX du signal audio est traité par un module 9 qui extrait des amplitudes spectrales associées aux harmoniques du signal correspondant aux multiples de la fréquence fondamentale estimée F_0 . Ces amplitudes sont ensuite interpolées par un module 10 afin d'obtenir une

enveloppe supérieure comprimée notée LX_{sup}.

Il est à noter que la compression spectrale pourrait de façon équivalente être effectuée après la détermination des amplitudes associées aux harmoniques. Elle pourrait également être effectuée après l'interpolation, ce qui ne ferait que modifier la forme des fonctions d'interpolation.

Le module 9 d'extraction des maxima tient compte de l'éventuelle variation de la fréquence fondamentale sur la trame d'analyse, des erreurs que peut commettre le détecteur 4, ainsi que des imprécisions liées au caractère discret de l'échantillonnage en fréquence. Pour cela, la recherche des amplitudes des pics spectraux ne consiste pas simplement à prendre les valeurs LX(i) correspondant aux index i tels que $i.F_0/2N$ soit la fréquence la plus proche d'une harmonique de fréquence $k.F_0$ ($k \geq 1$). L'amplitude spectrale retenue pour une harmonique d'ordre k est un maximum local du module du spectre au voisinage de la fréquence $k.F_0$ (cette amplitude est obtenue directement sous forme comprimée lorsque la compression spectrale 8 est effectuée avant l'extraction des maxima 9).

Les figures 4 et 5 montrent un exemple de forme du spectre comprimé LX, où on voit que les amplitudes maximales des pics harmoniques ne coïncident pas nécessairement avec les amplitudes correspondant aux multiples entiers de la fréquence fondamentale estimée F_0 . Les flancs des pics étant assez raides, une petite erreur de positionnement de la fréquence fondamentale F_0 , amplifiée par l'indice d'harmonique k, peut distordre fortement l'enveloppe supérieure estimée du spectre et provoquer une mauvaise modélisation de la structure formantique du signal. Par exemple, prendre directement l'amplitude spectrale pour la fréquence $3.F_0$ dans le cas des figures 4 et 5 produirait une erreur importante dans l'extraction de l'enveloppe supérieure au voisinage de l'harmonique d'ordre $k = 3$, alors qu'il s'agit d'une zone énergétiquement importante dans l'exemple dessiné. En effectuant l'interpolation à partir du véritable maximum, on évite ce genre d'erreur d'estimation de l'enveloppe supérieure.

Dans l'exemple représenté sur la figure 4, l'interpolation est effectuée entre des points dont l'abscisse est la fréquence correspondant au maximum de l'amplitude d'un pic spectral, et dont l'ordonnée est ce maximum, avant ou après compression.

L'interpolation effectuée pour calculer l'enveloppe supérieure LX_sup est une simple interpolation linéaire. Bien entendu une autre forme d'interpolation pourrait être utilisée (par exemple polynomiale ou spline).

5 Dans la variante préférée représentée sur la figure 5, l'interpolation est effectuée entre des points dont l'abscisse est une fréquence $k.F_0$ multiple de la fréquence fondamentale (en fait la fréquence la plus proche dans le spectre discret) et dont l'ordonnée est l'amplitude maximale, avant ou après compression, du spectre au voisinage de cette fréquence multiple.

10 En comparant les figures 4 et 5, on peut voir que le mode d'extraction selon la figure 5, qui repositionne les pics sur les fréquences harmoniques, conduit à une meilleure précision sur l'amplitude des pics que le décodeur attribuera aux fréquences multiples de la fréquence fondamentale. Il peut se produire un léger déplacement en fréquence de la position de ces pics, ce qui n'est pas perceptuellement très important et n'est d'ailleurs pas évité non plus
15 dans le cas de la figure 4. Dans le cas de la figure 4, les points d'ancrage pour l'interpolation sont confondus avec les sommets des pics harmoniques. Dans le cas de la figure 5, on impose que ces points d'ancrage se trouvent précisément aux fréquences multiples de la fréquence fondamentale, leurs amplitudes correspondant à celles des pics.

20 L'intervalle de recherche du maximum d'amplitude associé à une harmonique de rang k est centré sur l'index i de la fréquence de la TFR la plus proche de $k.F_0$, c'est-à-dire $i = \left\lfloor 2Nk \frac{F_0}{F_e} + \frac{1}{2} \right\rfloor$, où $\lfloor a \rfloor$ désigne l'entier égal ou immédiatement inférieur au nombre a . La largeur de cet intervalle de recherche dépend de la fréquence d'échantillonnage F_e , de la taille $2N$ de la TFR et de la
25 gamme de variation possible de la fréquence fondamentale. Cette largeur est typiquement de l'ordre d'une dizaine de fréquences avec les exemples de valeurs précédemment considérés. On peut la rendre réglable en fonction de la valeur F_0 de la fréquence fondamentale et du numéro k de l'harmonique.

30 Afin d'améliorer la résolution dans les basses fréquences et donc de représenter plus fidèlement les amplitudes des harmoniques dans cette zone, une distorsion non-linéaire de l'échelle des fréquences est opérée sur l'enveloppe supérieure comprimée par un module 12 avant que le module 13 effectue la transformée de Fourier rapide inverse (TFRI) fournissant les coefficients cepstraux cx_sup .

La distorsion non-linéaire permet de minimiser plus efficacement l'erreur de modélisation. Elle est par exemple effectuée selon une échelle de fréquences de type Mel ou Bark. Cette distorsion peut éventuellement dépendre de la fréquence fondamentale estimée F_0 . La figure 1 illustre le cas de l'échelle Mel. La relation entre les fréquences F du spectre linéaire, exprimées en hertz, et les fréquences F' de l'échelle Mel est la suivante :

$$F' = \frac{1000}{\log_{10}(2)} \times \log_{10} \left(1 + \frac{F}{1000} \right) \quad (6)$$

Afin de limiter le débit de transmission, une troncature des coefficients cepstraux cx_sup est effectuée. Le module de TFRI 13 a besoin de calculer seulement un vecteur cepstral de NCS coefficients cepstraux d'ordres 0 à NCS-1. A titre d'exemple, NCS peut être égal à 16.

Un post-filtrage dans le domaine cepstral, appelé post-liffrage, est appliqué par un module 15 à l'enveloppe supérieure comprimée LX_sup . Ce post-liffrage correspond à une manipulation des coefficients cepstraux cx_sup délivrés par le module de TRFI 13, qui correspond approximativement à un post-filtrage de la partie harmonique du signal par une fonction de transfert ayant la forme classique :

$$H(z) = \frac{(1 - \mu z^{-1}) A(z/\gamma_1)}{A(z/\gamma_2)} \quad (7)$$

où $A(z)$ est la fonction de transfert d'un filtre de prédiction linéaire du signal audio, γ_1 et γ_2 sont des coefficients compris entre 0 et 1, et μ est un coefficient de préaccentuation éventuellement nul. La relation entre le coefficient post-liffré d'ordre i , noté $c_p(i)$, et le coefficient cepstral correspondant $c(i) = cx_sup(i)$ délivré par le module 13 est alors :

$$\begin{aligned} c_p(0) &= c(0) \\ c_p(i) &= \left(1 + \gamma_2^i - \gamma_1^i \right) c(i) - \frac{\mu^i}{i} \quad \text{pour } i > 0 \end{aligned} \quad (8)$$

Le coefficient de préaccentuation optionnel μ peut être contrôlé en posant comme contrainte de préserver la valeur du coefficient cepstral $cx_sup(1)$ relatif à la pente. En effet, la valeur $c(1) = cx_sup(1)$ d'un bruit blanc filtré par le filtre de préaccentuation correspond au coefficient de préaccentuation. On peut ainsi choisir ce dernier de la façon suivante :

$$\mu = (\gamma_2 - \gamma_1) \cdot c(1).$$

Après le post-liffrage 15, un module de normalisation 16 modifie encore

les coefficients cepstraux en imposant la contrainte de modélisation exacte d'un point du spectre initial, qui est de préférence le point le plus énergétique parmi les maxima spectraux extraits par le module 9. En pratique, cette normalisation modifie seulement la valeur du coefficient $c_p(0)$.

- 5 Le module de normalisation 16 fonctionne de la façon suivante : il recalcule une valeur du spectre synthétisé à la fréquence du maximum indiqué par le module 9, par transformée de Fourier des coefficients cepstraux tronqués et post-liftrés, en tenant compte de la distorsion non-linéaire de l'axe des fréquences ; il détermine un gain de normalisation g_N par la différence
- 10 logarithmique entre la valeur du maximum fournie par le module 9 et cette valeur recalculée ; et il ajoute le gain g_N au coefficient cepstral post-liftré $c_p(0)$.

Cette normalisation peut être vue comme faisant partie du post-liftrage.

- Les coefficients cepstraux post-liftrés et normalisés font l'objet d'une quantification par un module 18 qui transmet des index de quantification
- 15 correspondants icxs au multiplexeur de sortie 6 du codeur.

- Le module 18 peut fonctionner par quantification vectorielle à partir de vecteurs cepstraux formés de coefficients post-liftrés et normalisés, notés ici $cx[n]$ pour la trame de signal de rang n . A titre d'exemple, le vecteur cepstral $cx[n]$ de $NCS = 16$ coefficients cepstraux $cx[n,0], cx[n,1], \dots, cx[n,NCS-1]$ est
- 20 distribué en quatre sous-vecteurs cepstraux contenant chacun quatre coefficients d'ordres consécutifs. Le vecteur cepstral $cx[n]$ peut être traité par les moyens représentés sur la figure 6, faisant partie du module de quantification 18. Ces moyens mettent en œuvre, pour chaque composante $cx[n,i]$, un prédicteur de la forme :

25
$$cx_p[n,i] = (1 - \alpha(i)).rcx[n,i] + \alpha(i).rcx[n-1,i] \quad (9)$$

- où $rcx[n]$ désigne un vecteur résiduel de prédiction pour la trame de rang n dont les composantes sont respectivement notées $rcx[n,0], rcx[n,1], \dots, rcx[n,NCS-1]$, et $\alpha(i)$ désigne un coefficient de prédiction choisi pour être représentatif d'une corrélation inter-trame supposée. Après quantification des
- 30 résidus, ce vecteur résiduel est défini par :

$$rcx[n,i] = \frac{cx[n,i] - \alpha(i).rcx_q[n-1,i]}{2 - \alpha(i)} \quad (10)$$

où $rcx_q[n-1]$ désigne le vecteur résiduel quantifié pour la trame de rang $n-1$, dont les composantes sont respectivement notées $rcx_q[n,0], rcx_q[n,1], \dots$,

rcx_q[n,NCS-1].

Le numérateur de la relation (10) est obtenu par un soustracteur 20, dont les composantes du vecteur de sortie sont divisées par les quantités $2^{-\alpha(i)}$ en 21. Aux fins de la quantification, le vecteur résiduel rcx[n] est subdivisé en quatre sous-vecteurs, correspondant à la subdivision en quatre sous-vecteurs cepstraux. Sur la base d'un dictionnaire obtenu par apprentissage préalable, l'unité 22 procède à la quantification vectorielle de chaque sous-vecteur du vecteur résiduel rcx[n]. Cette quantification peut consister, pour chaque sous-vecteur srcx[n], à sélectionner dans le dictionnaire le sous-vecteur quantifié srcx_q[n] qui minimise l'erreur quadratique $\|srcx[n] - srcx_q[n]\|^2$. L'ensemble icxs des index de quantification icx, correspondant aux adresses dans le ou les dictionnaires des sous-vecteurs résiduels quantifiés srcx_q[n], est fourni au multiplexeur de sortie 6.

L'unité 22 délivre également les valeurs des sous-vecteurs résiduels quantifiés, qui forment le vecteur rcx_q[n]. Celui-ci est retardé d'une trame en 23, et ses composantes sont multipliées par les coefficients $\alpha(i)$ en 24 pour fournir le vecteur à l'entrée négative du soustracteur 20. Ce dernier vecteur est d'autre part fourni à un additionneur 25, dont l'autre entrée reçoit un vecteur formé par les composantes du résidu quantifié rcx_q[n] respectivement multipliées par les quantités $1-\alpha(i)$ en 26. L'additionneur 25 délivre ainsi le vecteur cepstral quantifié cx_q[n] que récupérera le décodeur.

Le coefficient de prédiction $\alpha(i)$ peut être optimisé séparément pour chacun des coefficients cepstraux. Les dictionnaires de quantification peuvent aussi être optimisés séparément pour chacun quatre sous-vecteurs cepstraux. D'autre part, il est possible, de façon connue en soi, de normaliser les vecteurs cepstraux avant d'appliquer le schéma de prédiction/quantification, à partir de la variance des cepstres.

Il est à noter que le schéma ci-dessus de quantification des coefficients cepstraux peut n'être appliqué que pour certaines seulement des trames. Par exemple, on peut prévoir un second mode de quantification ainsi qu'un processus de sélection de celui des deux modes qui minimise un critère de moindres carrés avec les coefficients cepstraux à quantifier, et transmettre avec les index de quantification de la trame un bit indiquant lequel des deux modes a été sélectionné.

Les coefficients cepstraux quantifiés cx_sup_q = cx_q[n] fournis par

l'additionneur 25 sont adressés à un module 28 qui recalcule les amplitudes spectrales associées à une ou plusieurs des harmoniques de la fréquence fondamentale F_0 (figure 1). Ces amplitudes spectrales sont par exemple calculées sous forme comprimée, en appliquant la transformée de Fourier aux coefficients cepstraux quantifiés en tenant compte de la distorsion non-linéaire de l'échelle des fréquences utilisée dans la transformation cepstrale. Les amplitudes ainsi recalculées sont fournies à un module d'adaptation 29 qui les compare à des amplitudes de maxima déterminées par le module d'extraction 9.

Le module d'adaptation 29 contrôle le post-lifre 15 de façon à minimiser un écart de module entre le spectre du signal audio et les valeurs de module correspondantes calculées en 28. Cet écart de module peut être exprimé par une somme de valeurs absolues de différences d'amplitudes, comprimées ou non, correspondant à une ou plusieurs des fréquences harmoniques. Cette somme peut être pondérée en fonction des amplitudes spectrales associées à ces fréquences.

De façon optimale, l'écart de module pris en compte dans l'adaptation du post-lifrage tiendrait compte de toutes les harmoniques du spectre. Cependant, afin de réduire la complexité de l'optimisation, le module 28 peut ne resynthétiser les amplitudes spectrales que pour une ou plusieurs fréquences multiples de la fréquence fondamentale F_0 , sélectionnées sur la base de l'importance du module du spectre en valeur absolue. Le module d'adaptation 29 peut par exemple considérer les trois pics spectraux les plus intenses dans le calcul de l'écart de module à minimiser.

Dans une autre réalisation, le module d'adaptation 29 estime une courbe de masquage spectral du signal audio au moyen d'un modèle psychoacoustique, et les fréquences prises en compte dans le calcul de l'écart de module à minimiser sont sélectionnées sur la base de l'importance du module du spectre relativement à la courbe de masquage (on peut par exemple prendre les trois fréquences pour lesquelles le module du spectre dépasse le plus de la courbe de masquage). Différentes méthodes classiques sont utilisables pour calculer la courbe de masquage à partir du signal audio. On peut par exemple utiliser celle développée par J.D. Johnston (« Transform Coding of Audio Signals Using Perceptual Noise Criteria », IEEE Journal on Selected Area in Communications, Vol. 6, No. 2, février 1988).

Pour réaliser l'adaptation du post-lifrage, le module 29 peut utiliser un modèle d'identification de filtre. Une méthode plus simple consiste à prédéfinir un ensemble de jeux de paramètres de post-lifrage, c'est-à-dire un ensemble de couples γ_1, γ_2 dans le cas d'un post-lifrage selon les relations (8), à effectuer les opérations incombant aux modules 15, 16, 18 et 28 pour chacun de ces jeux de paramètres, et à retenir celui des jeux de paramètres qui conduit à l'écart de module minimal entre le spectre du signal et les valeurs recalculées. Les index de quantification fournis par le module 18 sont alors ceux qui se rapportent au meilleur jeu de paramètres.

Par un processus analogue à celui de l'extraction des coefficients cx_sup représentant l'enveloppe supérieure comprimée LX_sup du spectre du signal, le codeur détermine des coefficients cx_inf représentant une enveloppe inférieure comprimée LX_inf . Un module 30 extrait du spectre comprimé LX des amplitudes spectrales associées à des fréquences situées dans des zones du spectre intermédiaires par rapport aux fréquences multiples de la fréquence fondamentale estimée F_0 .

Dans l'exemple illustré par les figures 4 et 5, chaque amplitude associée à une fréquence située dans une zone intermédiaire entre deux harmoniques successives $k.F_0$ et $(k+1).F_0$ correspond simplement au module du spectre pour la fréquence $(k+1/2).F_0$ située au milieu de l'intervalle séparant les deux harmoniques. Dans une autre réalisation, cette amplitude pourrait être une moyenne du module du spectre sur une petite plage entourant cette fréquence $(k+1/2).F_0$.

Un module 31 procède à une interpolation, par exemple linéaire, des amplitudes spectrales associées aux fréquences situées dans les zones intermédiaires pour obtenir l'enveloppe inférieure comprimée LX_inf .

La transformation cepstrale appliquée à cette enveloppe inférieure comprimée LX_inf est effectuée suivant une échelle de fréquences résultant d'une distorsion non-linéaire appliquée par un module 32. Le module de TFRI 33 calcule un vecteur cepstral de NCI coefficients cepstraux cx_inf d'ordres 0 à NCI-1 représentant l'enveloppe inférieure. NCI est un nombre qui peut être sensiblement plus petit que NCS, par exemple $NCI = 4$.

La transformation non-linéaire de l'échelle des fréquences pour la transformation cepstrale de l'enveloppe inférieure peut être réalisée vers une échelle plus fine aux hautes fréquences qu'aux basses fréquences, ce qui

permet avantageusement de bien modéliser les composantes non-voisées du signal aux hautes fréquences. Toutefois, pour assurer une homogénéité de représentation entre l'enveloppe supérieure et l'enveloppe inférieure, on pourra préférer adopter dans le module 32 la même échelle que dans le module 12 (Mel dans l'exemple considéré).

Les coefficients cepstraux cx_inf représentant l'enveloppe inférieure comprimée sont quantifiés par un module 34, qui peut fonctionner de la même manière que le module 18 de quantification des coefficients cepstraux représentant l'enveloppe supérieure comprimée. Dans le cas considéré, où on se limite à $NCI = 4$ coefficients cepstraux pour l'enveloppe inférieure, le vecteur ainsi formé est soumis à une quantification vectorielle de résidu de prédiction, effectuée par des moyens identiques à ceux représentés sur la figure 6 mais sans subdivision en sous-vecteurs. L'index de quantification $icx = icxi$ déterminé par le quantificateur vectoriel 22 pour chaque trame relativement aux coefficients cx_inf est fourni au multiplexeur de sortie 6 du codeur.

Le codeur représenté sur la figure 1 ne comporte aucun dispositif particulier pour coder les phases du spectre aux harmoniques du signal audio.

En revanche, il comporte des moyens 36-40 pour coder une information temporelle liée à la phase de la composante non-harmonique représentée par l'enveloppe inférieure.

Un module 36 de décompression spectrale et un module 37 de TFRI forment une estimation temporelle de la trame de la composante non-harmonique. Le module 36 applique une fonction de décompression réciproque de la fonction de compression appliquée par le module 8 (c'est-à-dire une exponentielle ou une fonction puissance $1/\gamma$) à l'enveloppe inférieure comprimée LX_inf produite par le module d'interpolation 31. Ceci fournit le module de la trame estimée de la composante non-harmonique, dont la phase est prise égale à celle ϕ_X du spectre du signal X sur la trame. La transformée de Fourier inverse effectuée par le module 37 fournit la trame estimée de la composante non-harmonique.

Le module 38 subdivise cette trame estimée de la composante non-harmonique en plusieurs segments temporels. La trame délivrée par le module 37 se composant de $2N = 512$ échantillons pondérés comme illustré par la figure 3, le module 38 considère seulement les $N/2 = 128$ premiers échantillons et les $N/2 = 128$ derniers échantillons, et les subdivise par exemple en

huit segments de 32 échantillons consécutifs représentant chacun 4 ms de signal.

Pour chaque segment, le module 38 calcule l'énergie égale à la somme des carrés des échantillons, et forme un vecteur E1 formé de
5 huit composantes réelles positives égales aux huit énergies calculées. La plus grande de ces huit énergies, notée EM, est également déterminée pour être fournie, avec le vecteur E1, à un module de normalisation 39. Celui-ci divise chaque composante du vecteur E1 par EM, de sorte que le vecteur normalisé Emix est formé de huit composantes comprises entre 0 et 1. C'est ce vecteur
10 normalisé Emix, ou vecteur de pondération, qui est soumis à la quantification par le module 40. Celui-ci peut opérer une quantification vectorielle avec un dictionnaire déterminé lors d'un apprentissage préalable. L'index de quantification iEm est fourni par le module 40 au multiplexeur de sortie 6 du codeur.

15 La figure 7 montre une variante de réalisation des moyens employés par le codeur de la figure 1 pour déterminer le vecteur Emix de pondération énergétique de la trame de la composante non-harmonique. Les modules 36, 37 de décompression spectrale et de TFRI fonctionnent comme ceux qui portent les mêmes références sur la figure 1. Un module de sélection 42 est
20 ajouté pour déterminer la valeur du module du spectre soumis à la transformée de Fourier inverse 37. Sur la base de la fréquence fondamentale estimée F_0 , le module 42 identifie des régions harmoniques et des régions non-harmoniques du spectre du signal audio. Par exemple, une fréquence sera considérée comme appartenant à une région harmonique si elle se trouve dans un
25 intervalle de fréquences centré sur une harmonique $k.F_0$ et de largeur correspondant à une largeur de raie spectrale synthétisée, et à une région non-harmonique sinon. Dans les régions non-harmoniques, le signal complexe soumis à la TFRI 37 est égal à la valeur du spectre, c'est-à-dire que son module et sa phase correspondent aux valeurs $|X|$ et ϕ_X fournies par le module
30 de TFR 3. Dans les régions harmoniques, ce signal complexe a la même phase ϕ_X que le spectre et un module donné par l'enveloppe inférieure après décompression spectrale 36. Cette façon de procéder selon la figure 7 procure une modélisation plus précise des régions non-harmoniques.

Le décodeur représenté sur la figure 8 comprend un démultiplexeur
35 d'entrée 45 qui extrait du flux binaire Φ , issu d'un codeur selon la figure 1, les

index iF , $icxs$, $icxi$, iEm de quantification de la fréquence fondamentale F_0 , des coefficients cepstraux représentant l'enveloppe supérieure comprimée, des coefficients représentant l'enveloppe inférieure comprimée, et du vecteur de pondération $Emix$, et les distribue respectivement à des modules 46, 47, 48 et 49. Ces modules 46-49 comportent des dictionnaires de quantification semblables à ceux des modules 5, 18, 34 et 40 de la figure 1, afin de restituer les valeurs des paramètres quantifiés. Les modules 47 et 48 ont des dictionnaires pour former les résidus de prédiction quantifiés $rcx_q[n]$, et ils en déduisent les vecteurs cepstraux quantifiés $cx_q[n]$ avec des éléments identiques aux éléments 23-26 de la figure 6. Ces vecteurs cepstraux quantifiés $cx_q[n]$ fournissent les coefficients cepstraux cx_sup_q et cx_inf_q traités par le décodeur.

Un module 51 calcule la transformée de Fourier rapide des coefficients cepstraux cx_sup pour chaque trame de signal. L'échelle des fréquences du spectre comprimé qui en résulte est modifiée non-linéairement par un module 52 appliquant la transformation non-linéaire réciproque de celle du module 12 de la figure 1, et qui fournit l'estimation LX_sup de l'enveloppe supérieure comprimée. Une décompression spectrale de LX_sup , opérée par un module 53, fournit l'enveloppe supérieure X_sup comportant les valeurs estimées du module du spectre aux fréquences multiples de la fréquence fondamentale F_0 . Le module 54 synthétise l'estimation spectrale X_v de la composante harmonique du signal audio, par une somme de raies spectrales centrées sur les fréquences multiples de la fréquence fondamentale F_0 et dont les amplitudes (en module) sont celles données par l'enveloppe supérieure X_sup .

Bien que le flux numérique d'entrée Φ ne comporte pas d'informations spécifiques sur la phase du spectre du signal aux harmoniques de la fréquence fondamentale, le décodeur de la figure 8 est capable d'extraire de l'information sur cette phase à partir des coefficients cepstraux cx_sup_q représentant l'enveloppe supérieure comprimée. Cette information de phase est utilisée pour affecter une phase $\phi(k)$ à chacune des raies spectrales déterminées par le module 54 dans l'estimation de la composante harmonique du signal.

En première approximation, le signal de parole peut être considéré comme étant à phase minimale. D'autre part, il est connu que l'information de phase minimale peut se déduire facilement d'une modélisation cepstrale. Cette information de phase minimale est donc calculée pour chaque fréquence

harmonique. L'hypothèse de phase minimale signifie que l'énergie du signal synthétisé est localisée au début de chaque période de la fréquence fondamentale F_0 .

5 Pour être plus proche d'un signal de parole réel, on introduit un peu de dispersion au moyen d'un post-liffrage spécifique des cepstres lors de la synthèse de la phase. Avec ce post-liffrage, effectué par le module 55 de la figure 8, il est possible d'accentuer les résonances formantiques de l'enveloppe et donc de contrôler la dispersion des phases. Ce post-liffrage est par exemple de la forme (8).

10 Pour limiter les ruptures de phase, il est préférable de lisser les coefficients cepstraux post-liftrés, ce qui est effectué par le module 56. Le module 57 déduit des coefficients cepstraux post-liftrés et lissés la phase minimale affectée à chaque raie spectrale représentant un pic harmonique du spectre.

15 Les opérations effectuées par les modules 56, 57 de lissage et d'extraction de la phase minimale sont illustrées par l'organigramme de la figure 9. Le module 56 examine les variations des coefficients cepstraux pour appliquer un lissage moins important en présence de variations brusques qu'en présence de variations lentes. Pour cela, il effectue le lissage des coefficients cepstraux au moyen d'un facteur d'oubli λ_c choisi en fonction d'une
20 comparaison entre un seuil d_{th} et une distance d entre deux jeux successifs de coefficients cepstraux post-liftrés. Le seuil d_{th} est lui-même adapté en fonction des variations des coefficients cepstraux.

25 La première étape 60 consiste à calculer la distance d entre les deux vecteurs successifs relatifs aux trames $n-1$ et n . Ces vecteurs, notés ici $cxp[n-1]$ et $cxp[n]$, correspondent pour chaque trame à l'ensemble des NCS coefficients cepstraux post-liftrés représentant l'enveloppe supérieure comprimée. La distance utilisée peut notamment être la distance euclidienne entre les deux vecteurs ou encore une distance quadratique.

30 Deux lissages sont d'abord effectués, respectivement au moyen de facteurs d'oubli λ_{min} et λ_{max} , pour déterminer une distance minimale d_{min} et une distance maximale d_{max} . Le seuil d_{th} est ensuite déterminé à l'étape 70 comme étant situé entre les distances minimale et maximale d_{min} , d_{max} :
$$d_{th} = \beta \cdot d_{max} + (1-\beta) \cdot d_{min}$$
, le coefficient β étant par exemple égal à 0,5.

Dans l'exemple représenté, les facteurs d'oubli λ_{\min} et λ_{\max} sont eux-mêmes sélectionnés parmi deux valeurs distinctes, respectivement $\lambda_{\min1}$, $\lambda_{\min2}$ et $\lambda_{\max1}$, $\lambda_{\max2}$ comprises entre 0 et 1, les indices $\lambda_{\min1}$, $\lambda_{\max1}$ étant chacun sensiblement plus près de 0 que les indices $\lambda_{\min2}$, $\lambda_{\max2}$. Si $d > d_{\min}$ (test 61),

5 le facteur d'oubli λ_{\min} est égal à $\lambda_{\min1}$ (étape 62) ; sinon il est pris égal à $\lambda_{\min2}$ (étape 63). A l'étape 64, la distance minimale d_{\min} est prise égale à $\lambda_{\min} \cdot d_{\min} + (1 - \lambda_{\min}) \cdot d$. Si $d > d_{\max}$ (test 65), le facteur d'oubli λ_{\max} est égal à $\lambda_{\max1}$ (étape 66) ; sinon il est pris égal à $\lambda_{\max2}$ (étape 67). A l'étape 68, la distance minimale d_{\max} est prise égale à $\lambda_{\max} \cdot d_{\max} + (1 - \lambda_{\max}) \cdot d$.

10 Si la distance d entre les deux vecteurs cepstraux consécutifs est plus grande que le seuil d_{th} (test 71), on adopte pour le facteur d'oubli λ_c une valeur λ_{c1} relativement proche de 0 (étape 72). On considère dans ce cas que le signal correspondant est de type non stationnaire, de sorte qu'il n'y a pas lieu de conserver une grande mémoire des coefficients cepstraux antérieurs. Si

15 $d \leq d_{th}$, on adopte à l'étape 73 pour le facteur d'oubli λ_c une valeur λ_{c2} moins proche de 0 afin de lisser davantage les coefficients cepstraux. Le lissage est effectué à l'étape 74, où le vecteur $cxl[n]$ de coefficients lissés pour la trame courante n est déterminé par :

$$cxl[n] = \lambda_c \cdot cxl[n-1] + (1 - \lambda_c) \cdot cxp[n] \quad (11)$$

20 Le module 57 calcule ensuite les phases minimales $\phi(k)$ associées aux harmoniques $k \cdot F_0$. De façon connue, la phase minimale pour une harmonique d'ordre k est donnée par :

$$\phi(k) = -2 \cdot \sum_{m=1}^{NCS-1} cxl[n, m] \cdot \sin(2\pi m k F_0 / F_e) \quad (12)$$

où $cxl[n, m]$ désigne le coefficient cepstral lissé d'ordre m pour la trame n .

25 A l'étape 75, l'index d'harmonique k est initialisé à 1. Pour initialiser le calcul de la phase minimale affectée à l'harmonique k , la phase $\phi(k)$ et l'index cepstral m sont initialisés respectivement à 0 et 1 à l'étape 76. A l'étape 77, le module 57 ajoute à la phase $\phi(k)$ la quantité $-2 \cdot cxl[n, m] \cdot \sin(2\pi m k F_0 / F_e)$. L'index cepstral m est incrémenté à l'étape 78 et comparé à NCS à l'étape 79.

30 Les étapes 77 et 78 sont répétées tant que $m < NCS$. Quand $m = NCS$, le

calcul de la phase minimale est terminé pour l'harmonique k , et l'index k est incrémenté à l'étape 80. Le calcul de phases minimales 76-79 est renouvelé pour l'harmonique suivante tant que $k.F_0 < F_e/2$ (test 81).

Dans l'exemple de réalisation selon la figure 8, le module 54 tient
5 compte d'une phase constante sur la largeur de chaque raie spectrale, égale à la phase minimale $\varphi(k)$ fournie pour l'harmonique correspondante k par le module 57.

L'estimation X_v de la composante harmonique est synthétisée par
sommation de raies spectrales positionnées aux fréquences harmoniques de la
10 fréquence fondamentale F_0 . Lors de cette synthèse, on peut positionner les raies spectrales sur l'axe des fréquences avec une résolution supérieure à la résolution de la transformée de Fourier. Pour cela, on précalcule une fois pour toutes une raie spectrale de référence selon la résolution supérieure. Ce calcul
peut consister en une transformée de Fourier de la fenêtre d'analyse f_A avec
15 une taille de transformée de 16384 points, procurant une résolution de 0,5 Hz par point. La synthèse de chaque raie harmonique est alors effectuée par le module 54 en positionnant sur l'axe des fréquences la raie de référence à haute résolution, et en sous-échantillonnant cette raie spectrale de référence pour se ramener à la résolution de 16,625 Hz de la transformée de Fourier sur
20 512 points. Ceci permet de positionner avec précision la raie spectrale.

Pour la détermination de l'enveloppe inférieure, le module de TFR 85
du décodeur de la figure 8 reçoit les NCI coefficients cepstraux quantifiés
 cx_inf_q d'ordres 0 à $NCI - 1$, et il les complète avantageusement par les
NCS - NCI coefficients cepstraux cx_sup_q d'ordre NCI à NCS - 1
25 représentant l'enveloppe supérieure. En effet, on peut estimer en première approximation que les variations rapides de l'enveloppe inférieure comprimée sont bien reproduites par celles de l'enveloppe supérieure comprimée. Dans une autre réalisation, le module de TFR 85 pourrait ne considérer que les NCI
paramètres cepstraux cx_inf_q .

30 Le module 86 convertit l'échelle de fréquences de manière réciproque de la conversion opérée par le module 32 du codeur, afin de restituer l'estimation LX_inf de l'enveloppe inférieure comprimée, soumise au module de décompression spectrale 87. En sortie du module 87, le décodeur dispose d'une enveloppe inférieure X_inf comportant les valeurs du module du spectre

dans les vallées situées entre les pics harmoniques.

Cette enveloppe X_{inf} va moduler le spectre d'une trame de bruit dont la phase est traitée en fonction du vecteur de pondération quantifié E_{mix} extrait par le module 49. Un générateur 88 délivre une trame de bruit normalisé dont les segments de 4 ms sont pondérés dans un module 89 conformément aux composantes normalisées du vecteur E_{mix} fourni par le module 49 pour la trame courante. Ce bruit est un bruit blanc filtré passe-haut pour tenir compte du faible niveau qu'a en principe la composante non-voisée aux basses fréquences. A partir du bruit pondéré en énergie, le module 90 forme des trames de $2N = 512$ échantillons en appliquant la fenêtre d'analyse f_A , l'insertion de 256 échantillons à zéro et la permutation circulaire pour la compensation de phase conformément à ce qui a été expliqué en référence à la figure 3. La transformée de Fourier de la trame résultante est calculée par le module TFR 91.

L'estimation spectrale X_{uv} de la composante non-harmonique est déterminée par le module de synthèse spectrale 92 qui effectue une pondération fréquence par fréquence. Cette pondération consiste à multiplier chaque valeur spectrale complexe fournie par le module de TFR 91 par la valeur de l'enveloppe inférieure X_{inf} obtenue pour la même fréquence par le module de décompression spectrale 87.

Les estimations spectrales X_v , X_{uv} des composantes harmonique (voisée dans le cas d'un signal de parole) et non-harmonique (ou non-voisée) sont combinées par un module de mixage 95 contrôlé par un module 96 d'analyse du degré d'harmonicité (ou de voisement) du signal.

L'organisation de ces modules 95, 96 est illustrée par la figure 10. Le module d'analyse 96 comporte une unité 97 d'estimation d'un degré de voisement W dépendant de la fréquence, à partir duquel sont calculés quatre gains dépendant de la fréquence, à savoir deux gains g_v , g_{uv} contrôlant l'importance relative des composantes harmonique et non-harmonique dans le signal synthétisé, et deux gains g_{v_φ} , g_{uv_φ} utilisés pour bruite la phase de la composante harmonique.

Le degré de voisement $W(i)$ est une valeur à variation continue comprise entre 0 et 1 déterminée pour chaque index de fréquence i ($0 \leq i < N$) en fonction de l'enveloppe supérieure $X_{\text{sup}}(i)$ et de l'enveloppe inférieure $X_{\text{inf}}(i)$ obtenues pour cette fréquence i par les modules de décompression 53,

87. Le degré de voisement $W(i)$ est estimé par l'unité 97 pour chaque index de fréquence i correspondant à une harmonique de la fréquence fondamentale F_0 ,

à savoir $i = \left\lfloor 2Nk \frac{F_0}{F_e} + \frac{1}{2} \right\rfloor$ pour $k = 1, 2, \dots$, par une fonction croissante du

rapport entre l'enveloppe supérieure X_{sup} et l'enveloppe inférieure X_{inf} à cette fréquence, par exemple selon la formule :

$$W(i) = \min \left\{ 1, \frac{10 \cdot \log_{10} [X_{sup}(i)/X_{inf}(i)]}{V_{th}(F_0)} \right\} \quad (13)$$

Le seuil $V_{th}(F_0)$ correspond à la dynamique moyenne calculée sur un spectre synthétique purement voisé à la fréquence fondamentale. Il est avantageusement choisi dépendant de la fréquence fondamentale F_0 .

10 Le degré de voisement $W(i)$ pour une fréquence autre que les fréquences harmoniques est obtenu simplement comme étant égal à celui estimé pour l'harmonique la plus proche.

Le gain $g_v(i)$, qui dépend de la fréquence, est obtenu en appliquant une fonction non-linéaire au degré de voisement $W(i)$ (bloc 98). Cette fonction non-linéaire a par exemple la forme représentée sur la figure 11 :

$$g_v(i) = 0 \quad \text{si } 0 \leq W(i) \leq W1$$

$$g_v(i) = \frac{W(i) - W1}{W2 - W1} \quad \text{si } W1 < W(i) < W2 \quad (14)$$

$$g_v(i) = 1 \quad \text{si } W2 \leq W(i) \leq 1$$

les seuils $W1$, $W2$ étant tels que $0 < W1 < W2 < 1$. Le gain g_{uv} peut être calculé de manière semblable au gain g_v (la somme des deux gains g_v , g_{uv} étant constante, par exemple égale à 1), ou déduit simplement de celui-ci par la relation $g_{uv}(i) = 1 - g_v(i)$, comme schématisé par le soustracteur 99 sur la figure 10.

25 Il est intéressant de pouvoir bruite la phase de la composante harmonique du signal à une fréquence donnée si l'analyse du degré de voisement montre que le signal est plutôt de type non-harmonique à cette fréquence. Pour cela, la phase ϕ'_v de la composante harmonique mixée est le résultat d'une combinaison linéaire des phases ϕ_v , ϕ_{uv} des composantes harmonique et non-harmonique X_v , X_{uv} synthétisées par les modules 54, 92.

Les gains $g_{v_φ}$, $g_{uv_φ}$ respectivement appliqués à ces phases sont calculés à partir du degré de voisement W et pondérés également en fonction de l'index de fréquence i , étant donné que le bruitage de la phase n'est véritablement utile qu'au-delà d'une certaine fréquence.

5 Un premier gain $g_{v1_φ}$ est calculé en appliquant une fonction non-linéaire au degré de voisement $W(i)$, comme schématisé par le bloc 100 sur la figure 10. Cette fonction non-linéaire peut avoir la forme représentée sur la figure 12 :

$$\begin{aligned} g_{v1_φ}(i) &= G1 \quad \text{si } 0 \leq W(i) \leq W3 \\ 10 \quad g_{v1_φ}(i) &= G1 + (1 - G1) \frac{W(i) - W3}{W4 - W3} \quad \text{si } W3 < W(i) < W4 \\ g_{v1_φ}(i) &= 1 \quad \text{si } W4 \leq W(i) \leq 1 \end{aligned} \quad (15)$$

les seuils $W3$ et $W4$ étant tels que $0 < W3 < W4 < 1$, et le gain minimal $G1$ étant compris entre 0 et 1.

15 Un multiplieur 101 multiplie pour chaque fréquence d'index i le gain $g_{v1_φ}$ par un autre gain $g_{v2_φ}$ dépendant seulement de l'index de fréquence i , pour former le gain $g_{v_φ}(i)$. Le gain $g_{v2_φ}(i)$ dépend non-linéairement de l'index de fréquence i , par exemple comme indiqué sur la figure 13 :

$$\begin{aligned} g_{v2_φ}(i) &= 1 \quad \text{si } 0 \leq i \leq i1 \\ 20 \quad g_{v2_φ}(i) &= 1 - (1 - G2) \frac{i - i1}{i2 - i1} \quad \text{si } i1 < i < i2 \\ g_{v2_φ}(i) &= G2 \quad \text{si } i2 \leq i \leq 1 \end{aligned} \quad (16)$$

les index $i1$ et $i2$ étant tels que $0 < i1 < i2 \leq N$, et le gain minimal $G2$ étant compris entre 0 et 1. Le gain $g_{uv_φ}(i)$ peut être calculé simplement comme étant égal à $1 - g_{v_φ}(i) = 1 - g_{v1_φ}(i) \cdot g_{v2_φ}(i)$ (soustracteur 102 de la figure 10).

25 Le spectre complexe Y du signal synthétisé est produit par le module de mixage 95, qui réalise la relation de mixage suivante, pour $0 \leq i < N$:

$$Y(i) = g_v(i) \cdot |X_v(i)| \cdot \exp[j \phi_v(i)] + g_{uv}(i) \cdot X_{uv}(i) \quad (17)$$

avec
$$\phi_v(i) = g_{v_φ}(i) \cdot \phi_v(i) + g_{uv_φ}(i) \cdot \phi_{uv}(i) \quad (18)$$

où $\phi_v(i)$ désigne l'argument du nombre complexe $X_v(i)$ fourni par le module 54
30 pour la fréquence d'index i (bloc 104 de la figure 10), et $\phi_{uv}(i)$ désigne

l'argument du nombre complexe $X_{uv}(i)$ fourni par le module 92 (bloc 105 de la figure 10). Cette combinaison est réalisée par les multiplieurs 106-110 et les additionneurs 111-112 représentés sur la figure 10.

Le spectre mixé $Y(i)$ pour $0 \leq i < 2N$ (avec $Y(2N-1-i) = Y(i)$) est ensuite transformé dans le domaine temporel par le module de TFRI 115 (figure 8). On ne retient que les $N/2 = 128$ premiers et les $N/2 = 128$ derniers échantillons de la trame de $2N = 512$ échantillons produite par le module 115, et on applique la permutation circulaire inverse de celle illustrée par la figure 3 pour obtenir la trame synthétisée de $N = 256$ échantillons pondérés par la fenêtre d'analyse f_A .

Les trames successivement obtenues de cette manière sont finalement traitées par le module de synthèse temporelle 116 qui forme le signal audio décodé \hat{x} .

Le module de synthèse temporelle 116 effectue une somme à recouvrement de trames modifiées par rapport à celles successivement évaluées en sortie du module 115. La modification peut être vue en deux étapes illustrées respectivement par les figures 14 et 15.

La première étape (figure 14) consiste à multiplier chaque trame 2' délivrée par le module de TFRI 115 par une fenêtre $1/f_A$ inverse de la fenêtre d'analyse f_A employée par le module 1 du codeur. Les échantillons de la trame 2'' qui en résultent sont donc pondérés uniformément.

La seconde étape (figure 15) consiste à multiplier les échantillons de cette trame 2'' par une fenêtre de synthèse f_S vérifiant les propriétés suivantes :

$$f_S(N-L+i) + f_S(i) = A \quad \text{pour } 0 \leq i < L \quad (19)$$

$$f_S(i) = A \quad \text{pour } L \leq i < N-L \quad (20)$$

où A désigne une constante positive arbitraire, par exemple $A = 1$. La fenêtre de synthèse $f_S(i)$ croît progressivement de 0 à A pour i allant de 0 à L . C'est par exemple une demi-sinusoïde surélevée :

$$f_S(i) = \frac{A}{2} \cdot (1 - \cos[(i+1/2)\pi/L]) \quad \text{pour } 0 \leq i < L \quad (21)$$

Après avoir repondéré chaque trame 2'' par la fenêtre de synthèse f_S , le module 116 positionne les trames successives avec leurs décalages temporels de $M = 160$ échantillons et leurs recouvrements temporels de $L = 96$ échantillons, puis il effectue la somme des trames ainsi positionnées dans le temps. Du fait des propriétés (19) et (20) de la fenêtre de synthèse f_S , chaque

échantillon du signal audio décodé \hat{x} ainsi obtenu est affecté d'un poids global uniforme, égal à A. Ce poids global provient de la contribution d'une trame unique si l'échantillon a dans cette trame un rang i tel que $L \leq i < N - L$, et comporte les contributions sommées de deux trames successives si $0 \leq i < L$ ou $N - L \leq i < N$.

On peut ainsi effectuer la synthèse temporelle de façon simple même si, comme dans le cas considéré, le recouvrement L entre deux trames successives est plus petit que la moitié de la taille N de ces trames.

Les deux étapes exposées ci-dessus pour la modification des trames de signal peuvent être fusionnées en une seule étape. Il suffit de précalculer une fenêtre composée $f_C(i) = f_S(i)/f_A(i)$, et de multiplier simplement les trames $2'$ de $N = 256$ échantillons délivrées par le module 115 par la fenêtre composée f_C avant d'effectuer la sommation à recouvrement.

La figure 16 montre l'allure de la fenêtre composée f_C dans le cas où la fenêtre d'analyse f_A est une fenêtre de Hamming et la fenêtre de synthèse f_S a la forme donnée par les relations (19) à (21).

D'autres formes de la fenêtre de synthèse f_S vérifiant les relations (19) et (20) peuvent être employées. Dans la variante de la figure 17, c'est une fonction affine par morceaux définie par :

$$f_S(i) = A.i/L \quad \text{pour } 0 \leq i < L \quad (22)$$

Afin d'améliorer la qualité de codage du signal audio, le codeur de la figure 1 peut augmenter la cadence de formation et d'analyse des trames, afin de transmettre davantage de paramètres de quantification au décodeur. Dans la structure de trame représentée sur la figure 2, une trame de $N = 256$ échantillons (32 ms) est formée toutes les 20 ms. Ces trames de 256 échantillons pourraient être formées à une cadence supérieure, par exemple de 10 ms, deux trames successives ayant alors un décalage de $M/2 = 80$ échantillons et un recouvrement de 176 échantillons.

Dans ces conditions, on peut transmettre les jeux complets de paramètres de quantification iF , $icxs$, $icxi$, iEm pour seulement un sous-ensemble des trames, et transmettre pour les autres trames des paramètres permettant d'effectuer une interpolation adéquate au niveau du décodeur. Dans l'exemple envisagé ci-dessus, le sous-ensemble pour lequel des jeux de paramètres complets sont transmis peut être constitué par les trames de rang

entier n , dont la périodicité est de $M/F_e = 20$ ms, et les trames pour lesquelles une interpolation est effectuée peuvent être celles de rang demi-entier $n + 1/2$ qui sont décalées de 10 ms par rapport aux trames du sous-ensemble.

Dans la réalisation illustrée par la figure 18, les notations $cx_q[n-1]$ et $cx_q[n]$ désignent des vecteurs cepstraux quantifiés déterminés, pour deux trames successives de rang entier, par le module de quantification 18 et/ou par le module de quantification 34. Ces vecteurs comprennent par exemple quatre coefficients cepstraux consécutifs chacun. Ils pourraient également comprendre davantage de coefficients cepstraux.

Un module 120 effectue une interpolation de ces deux vecteurs cepstraux $cx_q[n-1]$ et $cx_q[n]$, afin d'estimer une valeur intermédiaire $cx_i[n-1/2]$. L'interpolation effectuée par le module 120 peut être une simple moyenne arithmétique des vecteurs $cx_q[n-1]$ et $cx_q[n]$. En variante, le module 120 pourrait appliquer une formule d'interpolation plus sophistiquée, par exemple polynomiale, en se fondant également sur les vecteurs cepstraux obtenus pour des trames antérieures à la trame $n-1$. D'autre part, si plus d'une trame interpolée est intercalée entre deux trames consécutives de rang entier, l'interpolation tient compte de la position relative de chaque trame interpolée.

A l'aide des moyens précédemment décrits, le codeur calcule également les coefficients cepstraux $cx[n-1/2]$ relatifs à la trame de rang demi-entier. Dans le cas de l'enveloppe supérieure, ces coefficients cepstraux sont ceux fournis par le module de TFRI 13 après post-lifrage 15 (par exemple avec les mêmes coefficients de post-lifrage que pour la trame précédente $n-1$) et normalisation 16. Dans le cas de l'enveloppe inférieure, les coefficients cepstraux $cx[n-1/2]$ sont ceux délivrés par le module de TFRI 33.

Un soustracteur 121 forme la différence $ecx[n-1/2]$ entre les coefficients cepstraux $cx[n-1/2]$ calculés pour la trame de rang demi-entier et les coefficients $cx_i[n-1/2]$ estimés par interpolation. Cette différence est fournie à un module de quantification 122 qui adresse des index de quantification $icx[n-1/2]$ au multiplexeur de sortie 6 du codeur. Le module 122 fonctionne par exemple par quantification vectorielle des erreurs d'interpolation $ecx[n-1/2]$ successivement déterminées pour les trames de rang demi-entier.

Cette quantification de l'erreur d'interpolation peut être effectuée par le codeur pour chacun des NCS + NCI coefficients cepstraux utilisés par le décodeur, ou seulement pour certains d'entre eux, typiquement ceux d'ordres

les plus petits.

Les moyens correspondants du décodeur sont illustrés par la figure 19. Le décodeur fonctionne essentiellement comme celui décrit en référence à la figure 8 pour déterminer les trames de signal de rang entier. Un module d'interpolation 124 identique au module 120 du codeur estime les coefficients intermédiaires $cx_i[n-1/2]$ à partir des coefficients quantifiés $cx_q[n-1]$ et $cx_q[n]$ fournis par le module 47 et/ou le module 48 à partir des index $icxs$, $icxi$ extraits du flux Φ . Un module d'extraction de paramètres 125 reçoit l'index de quantification $icx[n-1/2]$ depuis le démultiplexeur d'entrée 45 du décodeur, et en déduit l'erreur d'interpolation quantifiée $ecx_q[n-1/2]$ à partir du même dictionnaire de quantification que celui utilisé par le module 122 du codeur. Un additionneur 126 fait la somme des vecteurs cepstraux $cx_i[n-1/2]$ et $ecx_q[n-1/2]$ afin de fournir les coefficients cepstraux $cx[n-1/2]$ qui seront utilisés par le décodeur (modules 51-57, 95, 96, 115 et/ou modules 85-87, 92, 95, 96, 115) pour former la trame interpolée de rang $n-1/2$.

Si certains seulement des coefficients cepstraux ont fait l'objet d'une quantification d'erreur d'interpolation, les autres sont déterminés par le décodeur par une interpolation simple, sans correction.

Le décodeur peut également interpoler les autres paramètres F_0 , $Emix$ utilisés pour synthétiser les trames de signal. La fréquence fondamentale F_0 peut être interpolée linéairement, soit dans le domaine temporel, soit (de préférence) directement dans le domaine fréquentiel. Pour l'interpolation éventuelle du vecteur de pondération énergétique $Emix$, il convient d'effectuer l'interpolation après dénormalisation et en tenant compte bien entendu des décalages temporels entre trames.

Il est à noter qu'il est particulièrement avantageux, pour interpoler la représentation des enveloppes spectrales, d'effectuer cette interpolation dans le domaine cepstral. Contrairement à une interpolation effectuée sur d'autres paramètres, tels que les coefficients LSP (« Line Spectrum Pairs »), l'interpolation linéaire des coefficients cepstraux correspond à l'interpolation linéaire des amplitudes spectrales comprimées.

Dans la variante représentée sur la figure 20, le codeur utilise les vecteurs cepstraux $cx_q[n]$, $cx_q[n-1]$, ..., $cx_q[n-r]$ et $cx_q[n-1/2]$ calculés pour les dernières trames passées ($r \geq 1$) pour identifier un filtre interpolateur optimal qui, lorsqu'on lui soumet les vecteurs cepstraux quantifiés $cx_q[n-r]$,

..., $cx_q[n]$ relatifs aux trames de rang entier, délivre un vecteur cepstral interpolé $cx_i[n-1/2]$ qui présente une distance minimale avec le vecteur $cx[n-1/2]$ calculé pour la dernière trame de rang demi-entier.

Dans l'exemple représenté sur la figure 20, ce filtre interpolateur 128 est présent dans le codeur, et un soustracteur 129 retranche sa sortie $cx_i[n-1/2]$ du vecteur cepstral calculé $cx[n-1/2]$. Un module de minimisation 130 détermine le jeu de paramètres $\{P\}$ du filtre interpolateur 128, pour lequel l'erreur d'interpolation $ecx[n-1/2]$ délivrée par le soustracteur 129 présente une norme minimale. Ce jeu de paramètres $\{P\}$ est adressé à un module de quantification 131 qui fournit un index de quantification correspondant iP au multiplexeur de sortie 6 du codeur.

En fonction du débit alloué dans le flux Φ aux index de quantification des paramètres $\{P\}$ définissant le filtre interpolateur optimal 128, on pourra adopter une quantification plus ou moins fine de ces paramètres, ou une forme plus ou moins élaborée du filtre interpolateur, ou encore prévoir plusieurs filtres interpolateurs quantifiés de manière distincte pour différents vecteurs de coefficients cepstraux.

Dans une réalisation simple, le filtre interpolateur 128 est linéaire, avec $r = 1$:

$$cx_i[n-1/2] = \rho \cdot cx_q[n-1] + (1-\rho) \cdot cx_q[n] \quad (23)$$

et le jeu de paramètres $\{P\}$ se limite au coefficient ρ compris entre 0 et 1.

A partir des index iP de quantification des paramètres $\{P\}$ obtenus dans le flux binaire ϕ , le décodeur reconstruit le filtre interpolateur 128 (aux erreurs de quantification près), et traite les vecteurs spectraux $cx_q[n-r]$, ..., $cx_q[n]$ afin d'estimer les coefficients cepstraux $cx[n-1/2]$ utilisés pour synthétiser les trames de rang demi-entier.

De façon générale, le décodeur peut utiliser une méthode d'interpolation simple (sans transmission de paramètres de la part du codeur pour les trames de rang demi-entier), une méthode d'interpolation avec prise en compte d'une erreur d'interpolation quantifiée (selon les figures 17 et 18), ou une méthode d'interpolation avec un filtre interpolateur optimal (selon la figure 19) pour évaluer les trames de rang demi-entier en plus des trames de rang entier évaluées directement comme expliqué en référence aux figures 8 à 13. Le module 116 de synthèse temporelle peut alors combiner l'ensemble de ces trames évaluées pour former le signal synthétisé \hat{x} de la manière expliquée ci-

après en référence aux figures 14, 21 et 22.

Comme dans la méthode de synthèse temporelle précédemment décrite, le module 116 effectue une somme à recouvrement de trames modifiées par rapport à celles successivement évaluées en sortie du module 115, et cette modification peut être vue en deux étapes dont la première est identique à celle précédemment décrite en référence à la figure 14 (diviser les échantillons de la trame 2' par la fenêtre d'analyse f_A).

La seconde étape (figure 21) consiste à multiplier les échantillons de la trame renormalisée 2'' par une fenêtre de synthèse f'_S vérifiant les propriétés suivantes :

$$f'_S(i) = 0 \quad \text{pour } 0 \leq i < N/2 - M/p \quad \text{et} \quad N/2 + M/p \leq i < N \quad (24)$$

$$f'_S(i) + f'_S(i + M/p) = A \quad \text{pour } N/2 - M/p \leq i < N/2 \quad (25)$$

où A désigne une constante positive arbitraire, par exemple $A = 1$, et p est l'entier tel que le décalage temporel entre les trames successives (calculées directement et interpolées) soit de M/p échantillons, soit $p = 2$ dans l'exemple décrit. La fenêtre de synthèse $f'_S(i)$ croît progressivement pour i allant de $N/2 - M/p$ à $N/2$. C'est par exemple une sinusoïde surélevée sur l'intervalle $N/2 - M/p \leq i < N/2 + M/p$. En particulier, la fenêtre de synthèse f'_S peut être, sur cet intervalle, une fenêtre de Hamming (comme représenté sur la figure 21) ou une fenêtre de Hanning.

La figure 21 montre les trames successives 2'' repositionnées dans le temps par le module 116. Les hachures indiquent les portions éliminées des trames (fenêtre de synthèse à 0). On voit qu'en effectuant la somme à recouvrement des échantillons des trames successives, la propriété (25) assure une pondération homogène des échantillons du signal synthétisé.

Comme dans la méthode de synthèse illustrée par les figures 14 et 15, la procédure de pondération des trames obtenues par transformée de Fourier inverse des spectres Y peut être effectuée en une seule étape, avec une fenêtre composée $f'_C(i) = f'_S(i)/f_A(i)$. La figure 22 montre la forme de la fenêtre composée f'_C dans le cas où les fenêtres f_A et f'_S sont de type Hamming.

Comme la méthode de synthèse temporelle illustrée par les figures 14 à 17, celle illustrée par les figures 14, 21 et 22 permet de prendre en compte

un recouvrement L entre deux trames d'analyse (pour lesquelles l'analyse est effectuée de façon complète) plus petit que la moitié que la taille N de ces trames. De façon générale, cette dernière méthode est applicable lorsque les trames d'analyse successives présentent des décalages temporels mutuels M de plus de N/2 échantillons (même éventuellement de plus de N échantillons si un très bas débit est requis), l'interpolation conduisant à un ensemble de trames dont les décalages temporels mutuels sont de moins de N/2 échantillons.

Les trames interpolées peuvent faire l'objet d'une transmission réduite de paramètres de codage, comme décrit précédemment, mais cela n'est pas obligatoire. Ce mode de réalisation permet de conserver un intervalle M relativement grand entre deux trames d'analyse, et donc de limiter le débit de transmission requis, tout en limitant les discontinuités susceptibles d'apparaître en raison de la taille de cet intervalle par rapport aux échelles de temps typiques des variations des paramètres du signal audio, notamment les coefficients cepstraux et la fréquence fondamentale.

Les figures 23 à 25 montrent d'autres formes de réalisation des moyens employés pour traiter les coefficients cepstraux c_{x_sup} délivrés par le module de TFRI 13 de la figure 1, représentant l'enveloppe supérieure.

Dans les trois cas, les modules de post-lifrage 15, de normalisation 16, de quantification 18 et de calcul des amplitudes spectrales 28 sont essentiellement identiques à ceux précédemment décrits en référence à la figure 1. En outre, des modules de post-lifrage 140, de lissage 141 et d'extraction de phase minimale 142 sont prévus pour traiter les coefficients cepstraux post-liftrés et quantifiés $c_{x_sup_q}$ délivrés par le module de quantification 18. Ces modules 140-142 fonctionnent essentiellement comme les modules correspondants 55-57 du décodeur de la figure 8.

Dans la forme de réalisation montrée sur la figure 23, le module d'adaptation 144 accomplit une fonction semblable à celle du module 29 de la figure 1. Mais l'adaptation n'est pas réalisée sur la seule base du module du spectre. Le module 144 détermine le meilleur jeu de coefficients pour le post-lifre 15 en minimisant l'écart entre le spectre du signal audio, en module $|X|$ et en phase φ_X , et des valeurs complexes recalculées pour une ou plusieurs des harmoniques de la fréquence fondamentale. Les modules de ces dernières valeurs complexes sont données par le module de calcul 28, et leurs phases

correspondent aux phases minimales $\varphi(k)$ fournies par le module d'extraction 142. Pour réaliser l'adaptation, le module 144 peut prendre en compte toute distance appropriée dans le plan complexe, par exemple la distance euclidienne.

5 Ainsi, l'adaptation du post-lifre 15 par le module 144 tient compte de façon combinée d'aspects fréquentiels du signal, reflétés par le module du spectre, et des aspects temporels, reflétés par la phase du spectre.

 Comme représenté en pointillés sur la figure 23, le post-lifre 140 peut également être adaptatif, l'adaptation effectuée par le module 144 portant
10 conjointement sur les deux post-lifres 15, 140. Dans ce cas, le post-lifre 55 du décodeur (figure 8) est adapté, comme le post-lifre 140, en fonction de paramètres iLif que le module d'adaptation 144 fournit au multiplexeur 6 pour qu'il les inclue dans le flux numérique Φ . Typiquement, quelques jeux de coefficients γ_1 , γ_2 sont prévus pour les post-lifres 140 et 55, et le module 144
15 procède à un test exhaustif de ces différents jeux de coefficients pour retenir celui qui minimise l'écart dans le plan complexe.

 Dans l'exemple représenté sur la figure 24, le module d'adaptation 29 du post-lifre 15 est identique à celui de la figure 1. La figure 24 montre un module 145 d'estimation d'une courbe de masquage permettant au module 29
20 de sélectionner, pour la minimisation de l'écart en module, la ou les fréquences harmoniques qui dépassent le plus de la courbe de masquage calculée sur la base du spectre en module $|X|$, comme décrit précédemment.

 Le post-lifre 140 de la figure 24 est adapté séparément par un module 146 qui réalise la minimisation des écarts entre la phase φ_X du spectre du
25 signal et la phase minimale $\varphi(k)$ calculée par le module 142 pour une ou plusieurs des harmoniques. Là aussi, les harmoniques sélectionnées pour le calcul de l'écart de phase minimisé peuvent l'être en fonction de la courbe de masquage estimée par le module 145. Le module 146 fournit au multiplexeur de sortie 6 du codeur les paramètres iLif qui représentent le post-lifre optimal
30 140, pour qu'ils soient utilisés dans le post-filtre 55 du décodeur.

 Dans l'exemple illustré par la figure 25, le post-lifre 140 servant dans le calcul des phases minimales n'est pas adaptatif. Les phases minimales $\varphi(k)$ calculées par le module 142 pour les harmoniques de la fréquence fondamentale sont comparées aux phases φ_X du spectre du signal audio, et

l'écart de phase fait l'objet d'une quantification par un module 148. Les index de quantification correspondants $i\Delta\phi$ sont fournis par le module 148 au multiplexeur de sortie 6 du codeur.

Dans un décodeur (figure 26) correspondant à un codeur selon la figure 25, un module 149 exploite ces index de quantification $i\Delta\phi$ fournis par le démultiplexeur 45 pour obtenir les valeurs des écarts de phase quantifiés, qu'un additionneur 150 ajoute aux phases minimales $\phi(k)$ calculées par le module 57 (les post-filtres 140 et 55 étant identiques). Les phases fournies par l'additionneur 150 sont ensuite utilisées par le module 54 qui synthétise les raies spectrales de la composante harmonique X_v .

L'écart de phase quantifié par le module 148, et que les modules 149 et 150 du décodeur utilisent pour corriger les phases minimales $\phi(k)$, peut être de deux natures :

- il peut représenter, pour chaque fréquence d'index i correspondant à une harmonique d'ordre k de la fréquence fondamentale F_0 la différence entre la phase $\phi_X(i)$ du spectre du signal à la fréquence i et la phase minimale $\phi(k)$ calculée par le module 142 pour l'harmonique k ;
- alternativement ou cumulativement, cet écart de phase peut représenter la variation de la phase ϕ_X du spectre sur la largeur d'un ou plusieurs pics spectraux correspondant à des harmoniques du signal, cette variation étant relative à la phase minimale $\phi(k)$ affectée aux pics en question.

Dans les deux cas, le ou les pics pour lesquels l'écart de phase est quantifié peuvent être choisis en fonction de l'énergie spectrale représentée par l'enveloppe supérieure, qui est disponible au codeur et au décodeur, ce qui permet au décodeur de déterminer à quelle raie spectrale les écarts doivent être appliqués.

Dans le premier cas, les écarts de phase peuvent faire l'objet d'une quantification scalaire, ou vectorielle si on les regroupe pour plusieurs pics.

Dans le second cas, la variation de la phase ϕ_X autour de la phase minimale $\phi(k)$ sur la largeur d'un pic harmonique (déterminée par la largeur de la raie de référence utilisée par le module 54), peut être représentée simplement par la pente d'un segment linéaire sélectionné comme étant celui qui présente une distance quadratique minimale avec la courbe de variation de

phase du spectre sur la largeur de la raie, et éventuellement par un décalage à l'origine.

Ces pentes peuvent faire l'objet d'une quantification scalaire, ou vectorielle si on les regroupe pour plusieurs pics.

5 La quantification des variations de phase sur les pics harmoniques peut porter sur l'ensemble des fréquences harmoniques. Une autre possibilité est de quantifier plusieurs pentes obtenues chacune par moyennage des pentes aux harmoniques sur une ou plusieurs sous-bandes du spectre. Ce moyennage peut être pondéré pour prendre en compte les énergies relatives
10 aux différentes fréquences harmoniques, représentées par l'enveloppe supérieure.

 Le module 148 peut également modéliser la variation de phase sur la largeur d'un pic par une courbe plus complexe qu'un segment linéaire, par exemple un spline, dont les paramètres sont quantifiés pour être transmis au
15 décodeur.

 Une autre possibilité est d'effectuer un apprentissage préalable de modèles de phase aux harmoniques, représentatifs des variations de la phase sur la largeur des pics, observées sur un corpus de signaux de référence. Ces modèles sont rangés dans un dictionnaire mémorisé par les modules 148 et
20 149. Le module 148 du codeur détermine les index $i\Delta\phi$ correspondant aux adresses des modèles les plus proches des variations de phase au voisinage des pics harmoniques considérés, et le module 149 du décodeur récupère ces modèles pour la synthèse de la phase de la composante harmonique.

REVENDICATIONS

1. Procédé de codage d'un signal audio, dans lequel on estime une fréquence fondamentale (F_0) du signal audio, on détermine un spectre du signal audio par une transformée dans le domaine fréquentiel d'une trame du signal audio, et on inclut dans un flux numérique de sortie (Φ) des données de codage d'une composante harmonique du signal audio, comprenant des données représentatives d'amplitudes spectrales associées à des fréquences multiples de la fréquence fondamentale estimée, caractérisé en ce que lesdites données représentatives d'amplitudes spectrales associées à des fréquences multiples de la fréquence fondamentale (F_0) sont obtenues au moyen de coefficients cepstraux calculés en transformant dans le domaine cepstral une enveloppe supérieure comprimée (LX_sup) du spectre du signal audio, et en ce que les données de codage de la composante harmonique comprennent en outre, pour au moins une des fréquences multiples de la fréquence fondamentale estimée, des données ($i\Delta\phi$; $iLif$) relatives à la phase du spectre du signal audio au voisinage de ladite fréquence multiple.
2. Procédé selon la revendication 1, dans lequel l'enveloppe supérieure comprimée (LX_sup) est déterminée par interpolation desdites amplitudes spectrales associées aux fréquences multiples de la fréquence fondamentale (F_0) avec application d'une fonction de compression spectrale.
3. Procédé selon la revendication 1 ou 2, dans lequel on transforme les coefficients cepstraux par liftrage et lissage dans le domaine cepstral, et on calcule des phases minimales ($\phi(k)$) du signal audio à des fréquences multiples de la fréquence fondamentale (F_0) sur la base des coefficients cepstraux transformés ($cxl[n]$),
4. Procédé selon la revendication 3, dans lequel on quantifie un écart de phase entre le spectre du signal audio à au moins une fréquence multiple de la fréquence fondamentale (F_0) et la phase minimale calculée correspondante ($\phi(k)$), et on inclut des paramètres de quantification ($i\Delta\phi$) dudit écart de phase dans les données de codage de la composante harmonique.

5. Procédé selon la revendication 3, dans lequel on adapte ledit liftrage de façon à minimiser un écart de phase entre le spectre du signal audio à au moins une fréquence multiple de la fréquence fondamentale (F_0) et la phase minimale calculée correspondante ($\varphi(k)$), et on inclut des paramètres ($iLif$) représentatifs du liftrage adapté dans les données de codage de la composante harmonique.

6. Procédé selon la revendication 5, dans lequel l'écart de phase minimisé pour l'adaptation du liftrage se rapporte à au moins une fréquence multiple de la fréquence fondamentale (F_0), sélectionnée sur la base de l'importance du module du spectre en valeur absolue.

7. Procédé selon la revendication 5, dans lequel on estime une courbe de masquage spectral du signal audio au moyen d'un modèle psycho-acoustique, et l'écart de phase minimisé pour l'adaptation du liftrage se rapporte à au moins une fréquence multiple de la fréquence fondamentale (F_0), sélectionnée sur la base de l'importance du module du spectre relativement à la courbe de masquage.

8. Procédé selon l'une quelconque des revendications 3 à 7, dans lequel on examine les variations des coefficients cepstraux ($c_{xp}[n]$) pour appliquer un lissage moins important en présence de variations brusques qu'en présence de variations lentes.

9. Procédé selon la revendication 8, dans lequel on effectue le lissage des coefficients cepstraux au moyen d'un facteur d'oubli (λ_c) choisi en fonction d'une comparaison entre un seuil (d_{th}) et une distance (d) entre deux jeux successifs de coefficients cepstraux ($c_{xp}[n-1]$, $c_{xp}[n]$).

10. Procédé selon la revendication 9, dans lequel ledit seuil (d_{th}) est adapté en fonction des variations des coefficients cepstraux ($c_{xp}[n]$).

11. Procédé selon l'une quelconque des revendications 3 à 10, dans lequel le liftrage est de la forme $c_p(i) = [1 + \gamma_2^i - \gamma_1^i].c(i) - (\mu^i/i)$, où $c_p(i)$ et $c(i)$ désignent le coefficient cepstral d'ordre i respectivement avant et après liftrage,

γ_1 et γ_2 sont des coefficients compris entre 0 et 1 et μ est un coefficient de préaccentuation.

12. Procédé selon la revendication 11, dans lequel $\mu = (\gamma_2 - \gamma_1) \cdot c(1)$.

13. Procédé selon l'une quelconque des revendications 1 à 12, dans lequel on quantifie une variation de la phase du spectre du signal audio sur au moins un intervalle de fréquences entourant une fréquence multiple de la fréquence fondamentale (F_0), et on inclut des paramètres ($i\Delta\phi$) de quantification de ladite variation de phase dans les données de codage de la composante harmonique.

14. Procédé selon la revendication 13, dans lequel la variation de phase quantifiée est représentée par la pente de la phase du spectre à la fréquence multiple de la fréquence fondamentale (F_0).

15. Procédé selon la revendication 13, dans lequel la variation de phase quantifiée représente un modèle de phase mémorisé dans un dictionnaire déterminé par apprentissage.

16. Codeur audio, comprenant des moyens pour exécuter un procédé selon l'une quelconque des revendications 1 à 15.

17. Procédé de décodage d'un flux numérique d'entrée (Φ) représentant un signal audio codé, dans lequel on reçoit dans le flux numérique d'entrée des données de codage d'une fréquence fondamentale (F_0) du signal audio et des données de codage d'une composante harmonique comprenant des données de quantification de coefficients cepstraux représentant une enveloppe spectrale supérieure (X_{sup}) du signal audio, et on génère une estimation spectrale (X_v) de la composante harmonique dont le module est déterminé par ladite enveloppe spectrale supérieure aux fréquences multiples de la fréquence fondamentale, caractérisé en ce que, pour déterminer la phase (ϕ_v) de l'estimation spectrale de la composante harmonique (X_v), on transforme les coefficients cepstraux par liftrage et lissage dans le domaine cepstral, et on calcule des phases minimales ($\phi(k)$) du signal audio à des fréquences multiples

de la fréquence fondamentale sur la base des coefficients cepstraux transformés ($cxl[n]$).

18. Procédé selon la revendication 17, dans lequel, pour obtenir la phase (φ_v) de l'estimation spectrale (X_v) de la composante harmonique, on modifie au moins une phase minimale calculée ($\varphi(k)$) sur la base de données de quantification d'écart de phase ($i\Delta\varphi$) comprises dans les données de codage de la composante harmonique.

18. Procédé selon la revendication 18, dans lequel les données de quantification d'écart de phase ($i\Delta\varphi$) sont utilisées pour déterminer une variation de la phase (φ_v) de l'estimation spectrale (X_v) de la composante harmonique sur au moins un intervalle de fréquences entourant une fréquence multiple de la fréquence fondamentale (F_0).

20. Procédé selon la revendication 19, dans lequel la variation de phase déterminée sur un intervalle de fréquences est une variation linéaire représentée par un paramètre de quantification de pente.

21. Procédé selon la revendication 19, dans lequel la variation de phase déterminée sur un intervalle de fréquences représente un modèle de phase mémorisé dans un dictionnaire déterminé par apprentissage.

22. Procédé selon l'une quelconque des revendications 17 à 21, dans lequel on adapte ledit lissage en fonction de paramètres ($iLif$) compris dans les données de codage de la composante harmonique.

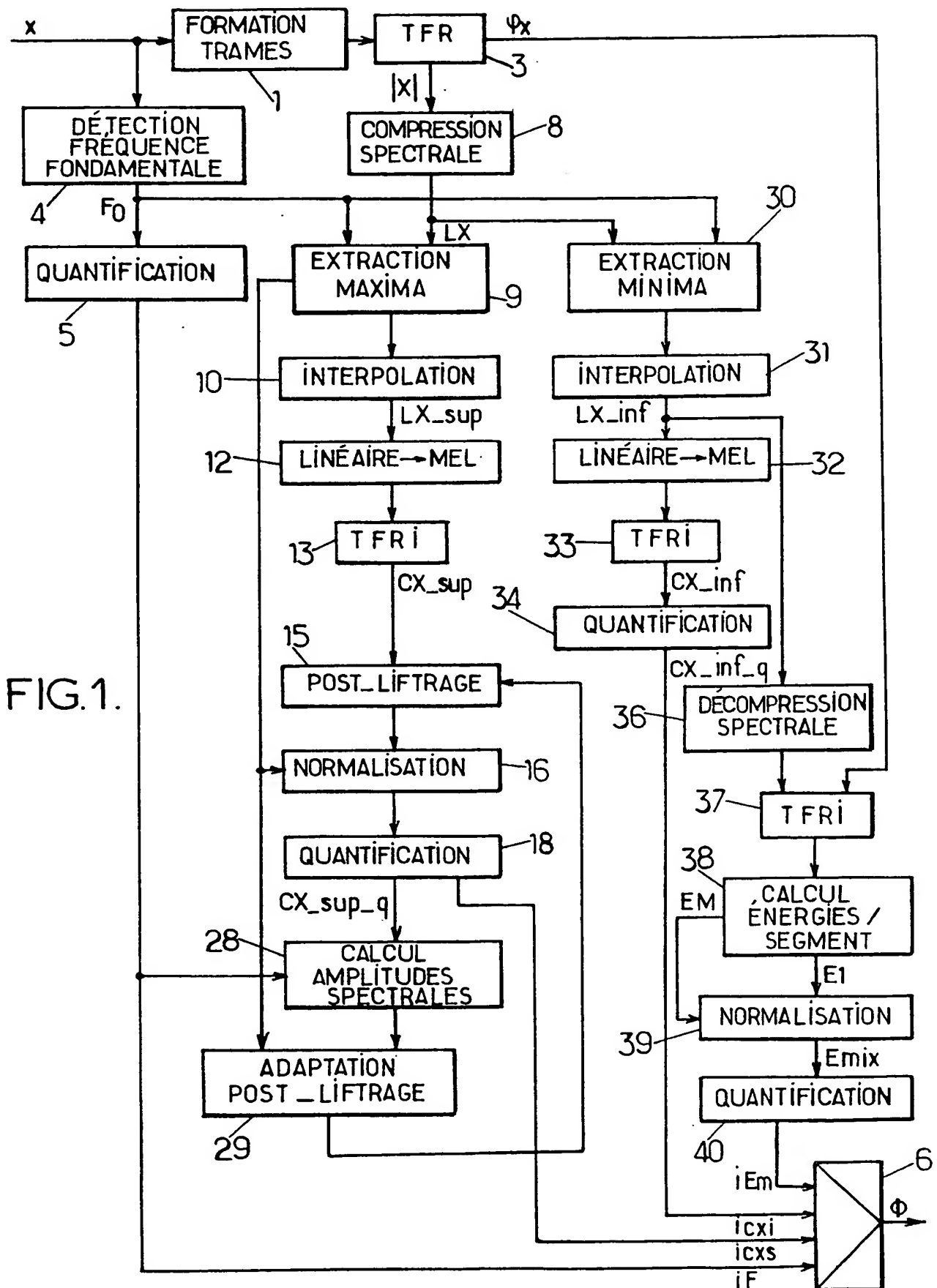
23. Procédé selon l'une quelconque des revendications 17 à 22, dans lequel on examine les variations des coefficients cepstraux ($cxp[n]$) pour appliquer un lissage moins important en présence de variations brusques qu'en présence de variations lentes.

24. Procédé selon la revendication 23, dans lequel on effectue le lissage des coefficients cepstraux ($cxp[n]$) au moyen d'un facteur d'oubli (λ_c) choisi en fonction d'une comparaison entre un seuil (d_{th}) et une distance (d) entre deux jeux successifs de coefficients cepstraux ($cxp[n-1]$, $cxp[n]$).

25. Procédé selon la revendication 24, dans lequel ledit seuil (d_{th}) est adapté en fonction des variations des coefficients cepstraux ($c_{xp}[n]$).
26. Procédé de décodage d'un flux numérique d'entrée (Φ) représentant un signal audio codé, dans lequel on reçoit dans le flux numérique d'entrée des données de codage d'une fréquence fondamentale (F_0) du signal audio et des données de codage d'une composante harmonique comprenant des données de quantification de coefficients cepstraux représentant une enveloppe spectrale supérieure (X_{sup}) du signal audio, et on génère une estimation spectrale (X_v) de la composante harmonique dont le module est déterminé par ladite enveloppe spectrale supérieure aux fréquences multiples de la fréquence fondamentale, caractérisé en ce qu'on détermine une variation de la phase (ϕ_v) de l'estimation spectrale de la composante harmonique sur au moins un intervalle de fréquences entourant une fréquence multiple de la fréquence fondamentale sur la base de paramètres de quantification ($i\Delta\phi$) inclus dans les données de codage de la composante harmonique.
27. Procédé selon la revendication 26, dans lequel la variation de phase déterminée sur un intervalle de fréquences est une variation linéaire représentée par un paramètre de quantification de pente.
28. Procédé selon la revendication 26, dans lequel la variation de phase déterminée sur un intervalle de fréquences représente un modèle de phase mémorisé dans un dictionnaire déterminé par apprentissage.
29. Décodeur audio, comprenant des moyens pour exécuter un procédé selon l'une quelconque des revendications 17 à 28.

THIS PAGE BLANK (USPTO

1/13



THIS PAGE BLANK (USPTO)

FIG. 2

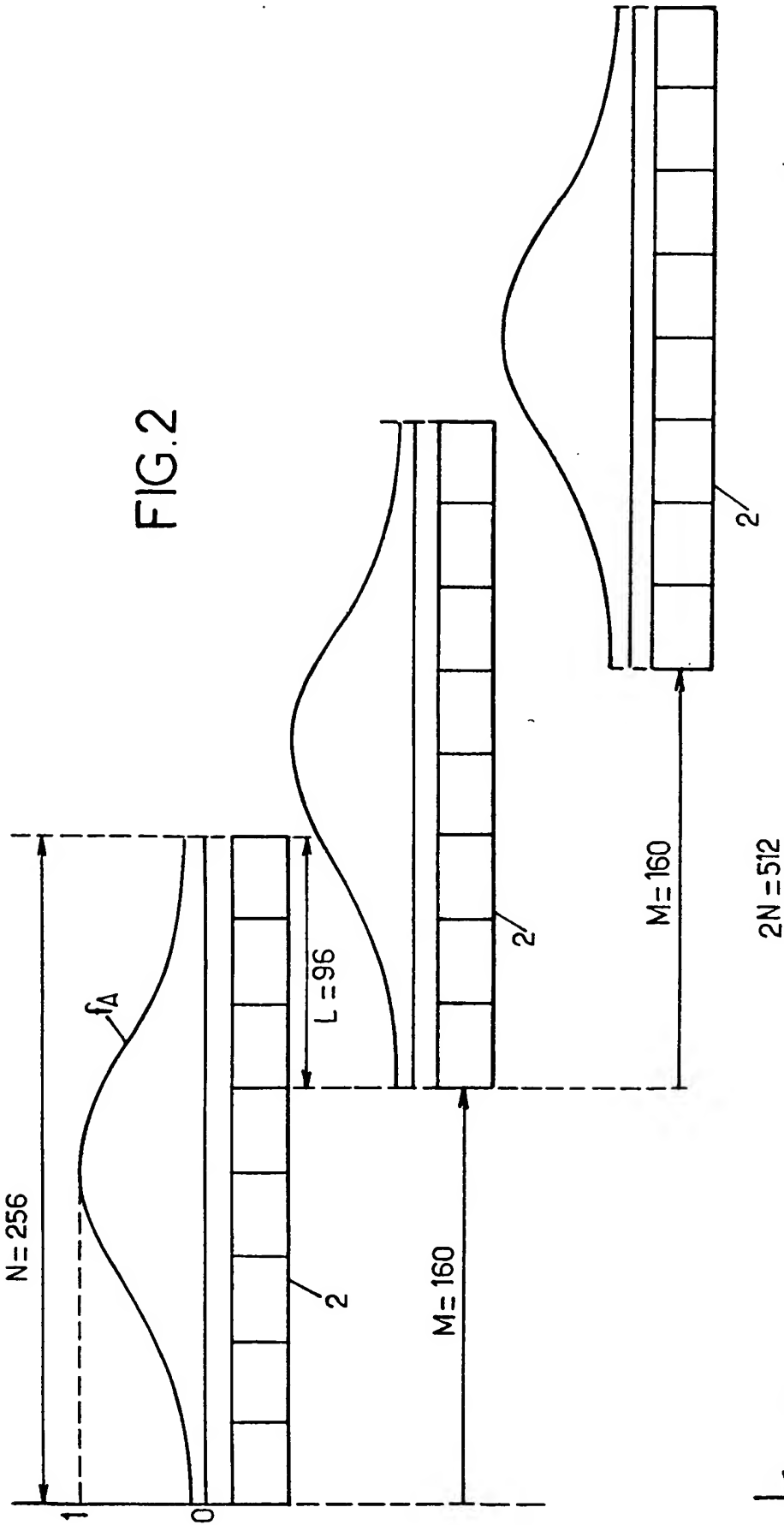
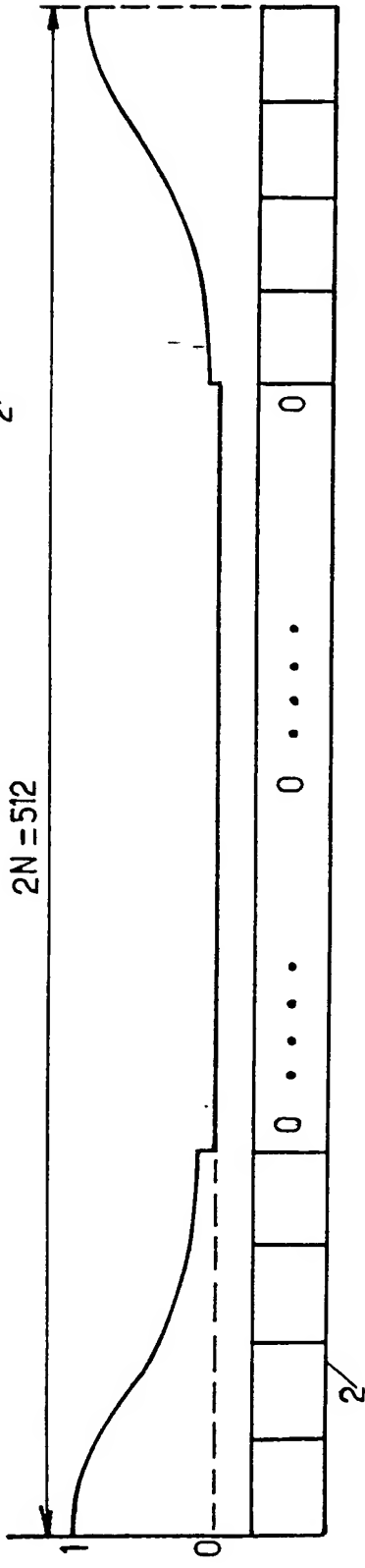


FIG. 3



THIS PAGE BLANK (USPTO)

3/13

FIG.4.

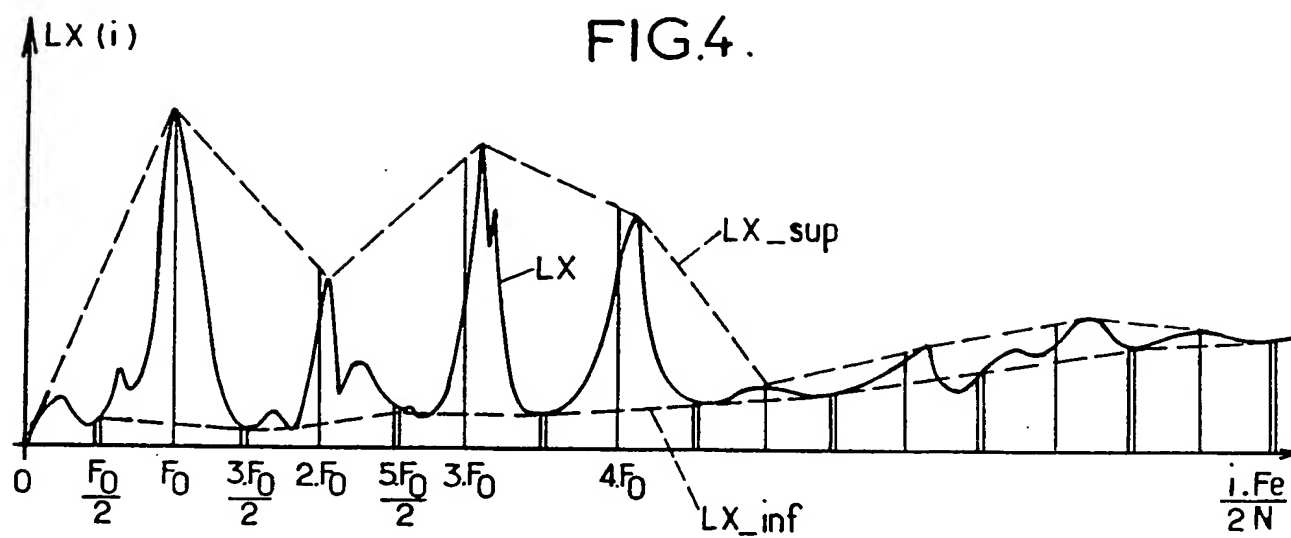
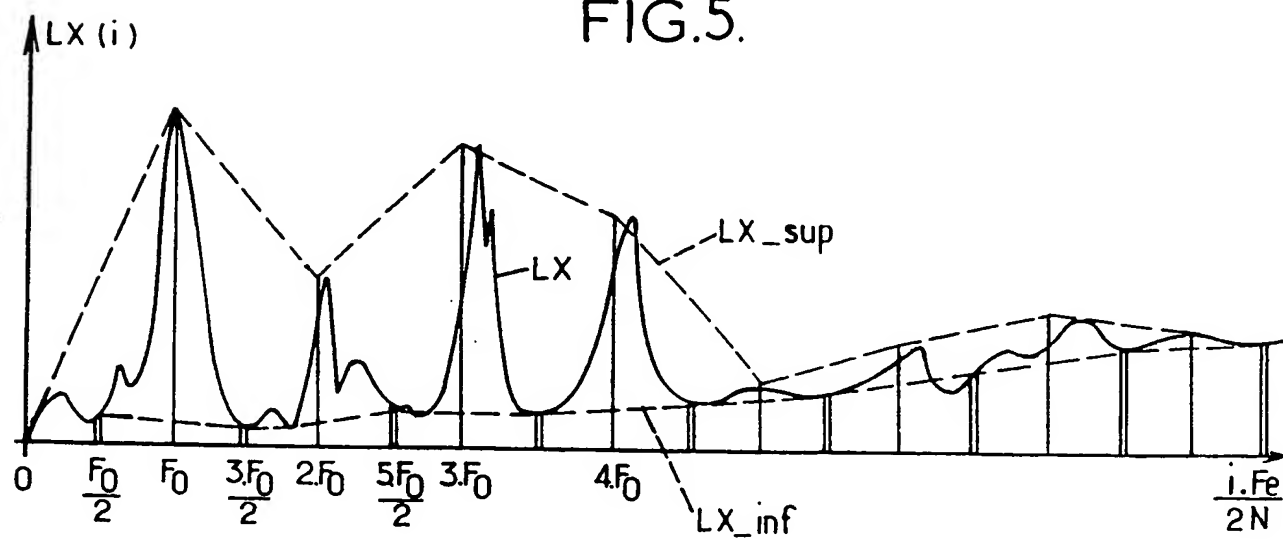


FIG.5.



THIS PAGE BLANK (USPTO)

4/13

FIG. 6.

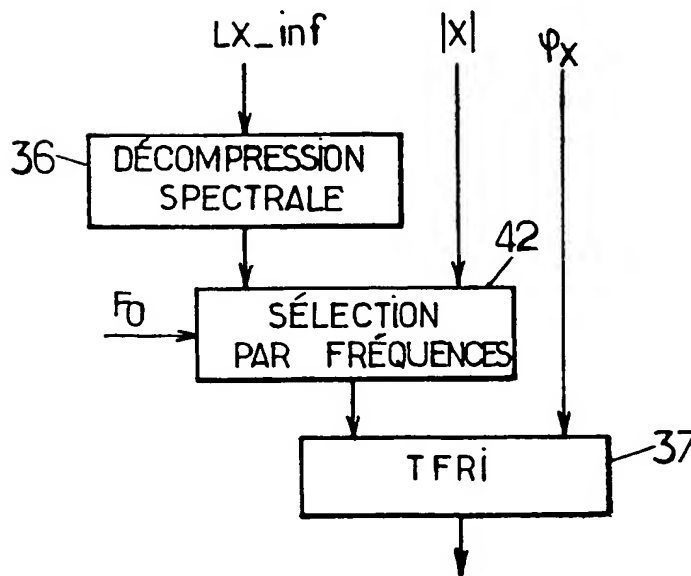
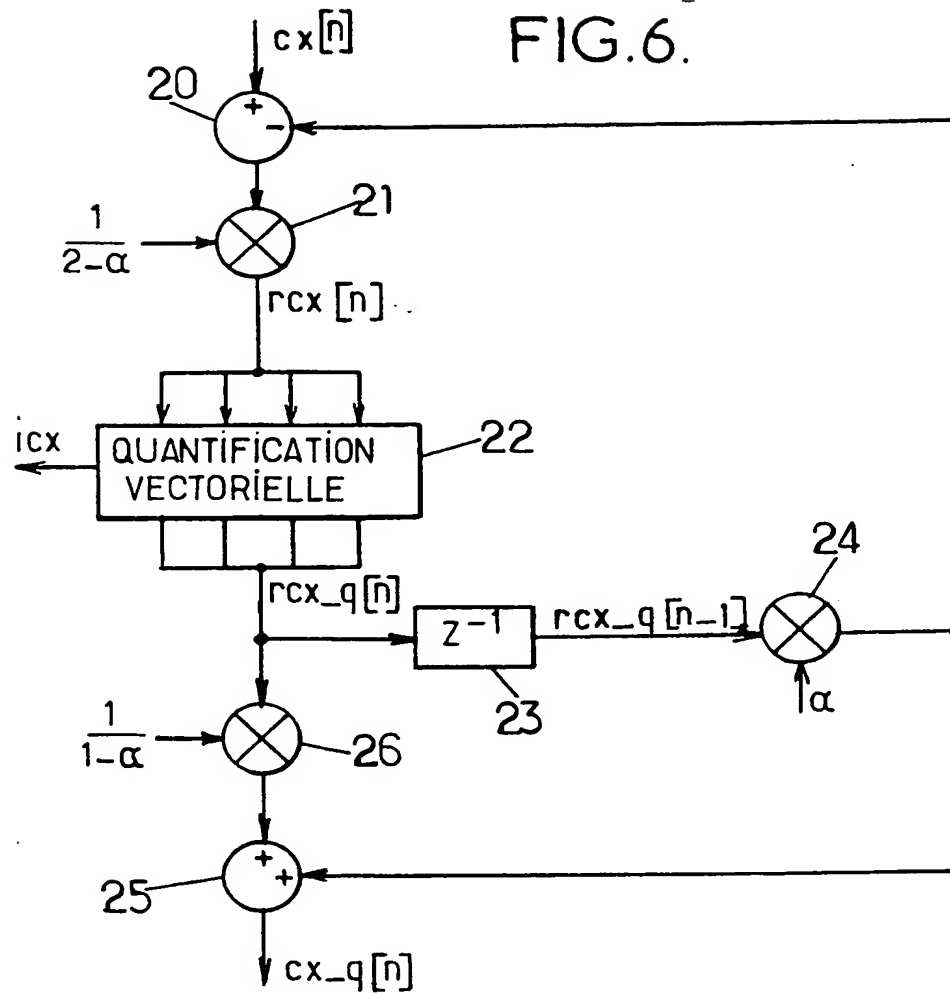


FIG. 7.

THIS PAGE BLANK (USPTO,

5/13

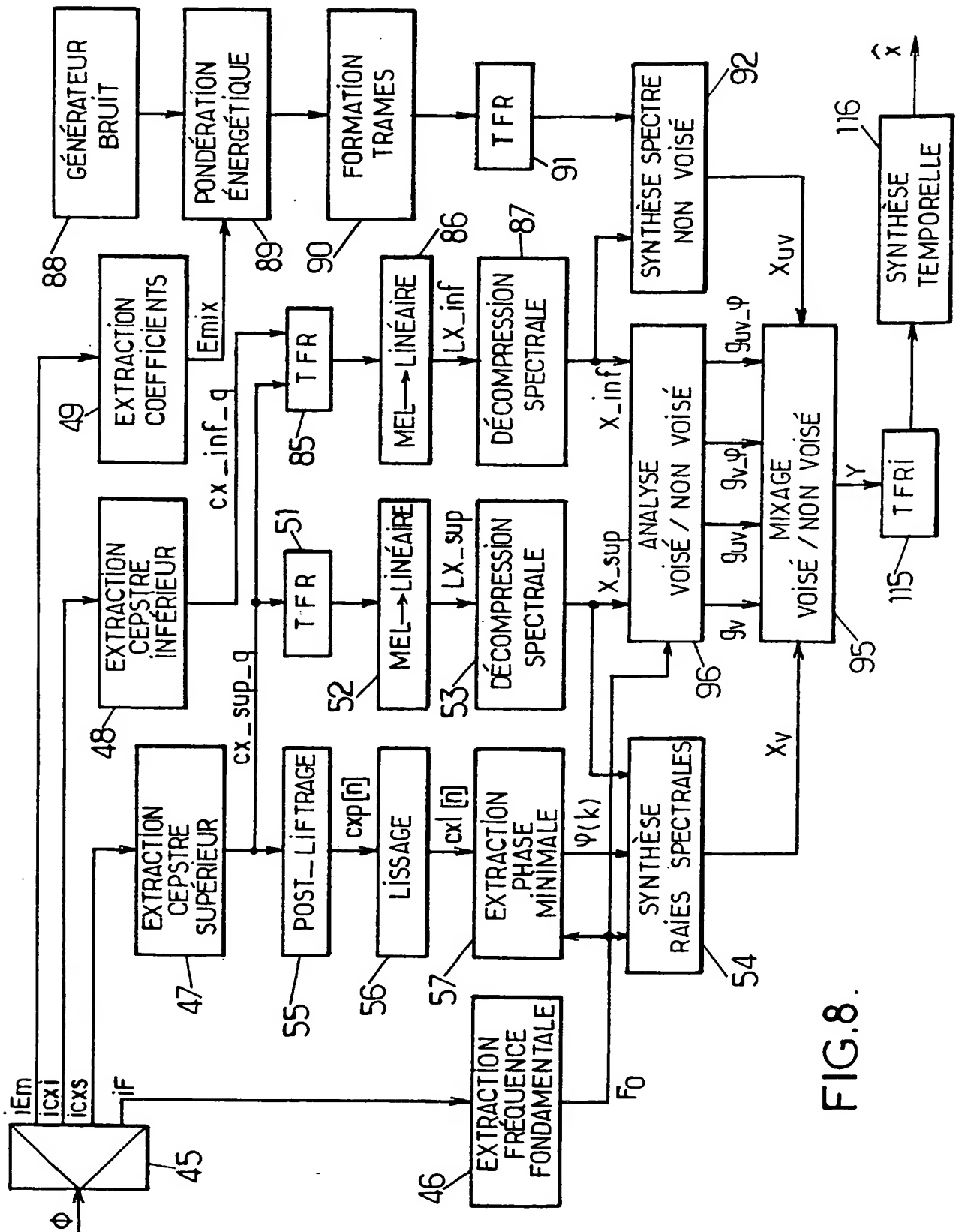
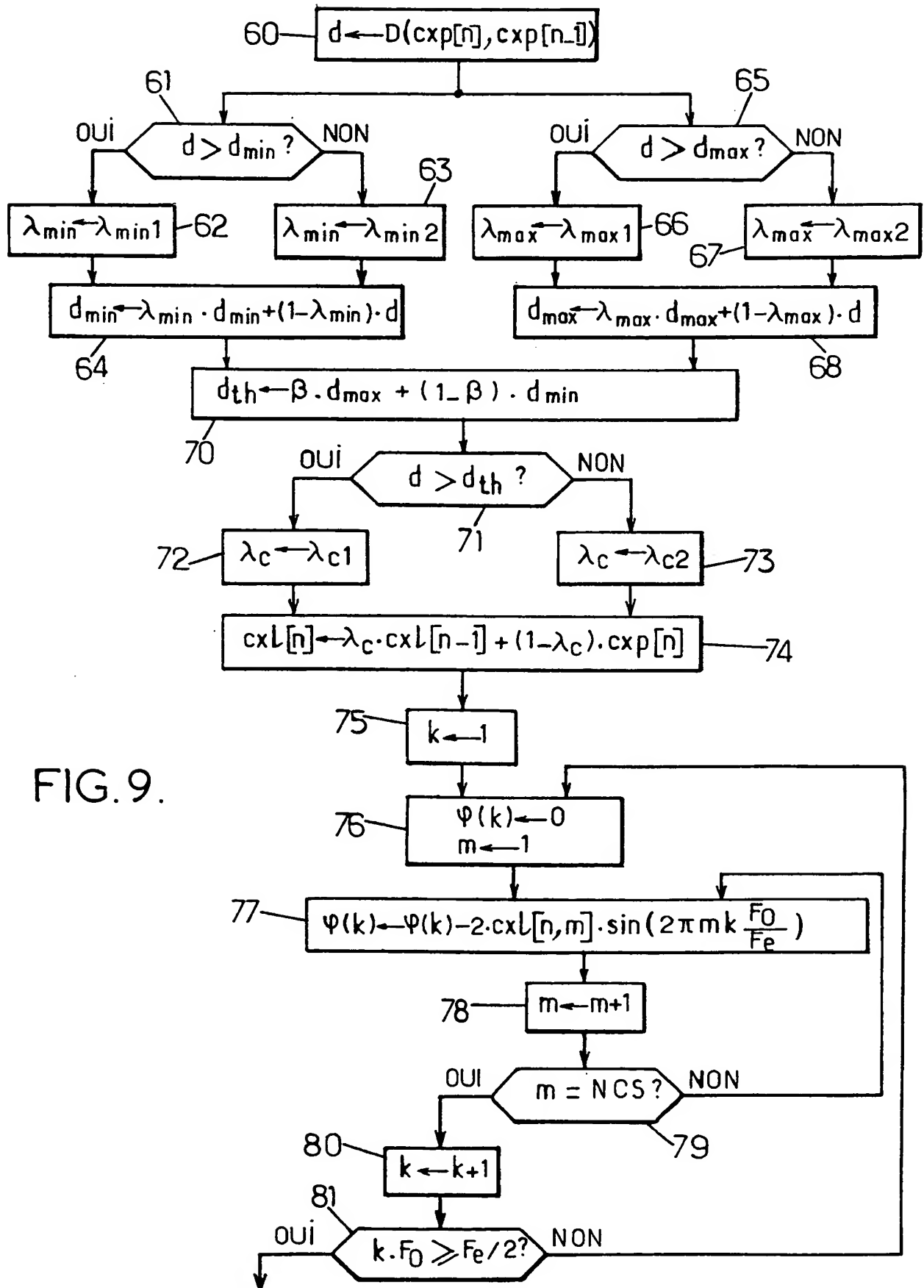


FIG. 8.

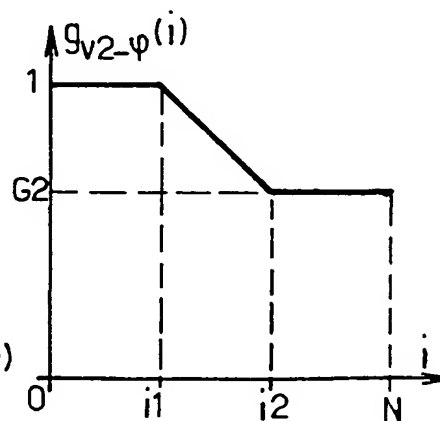
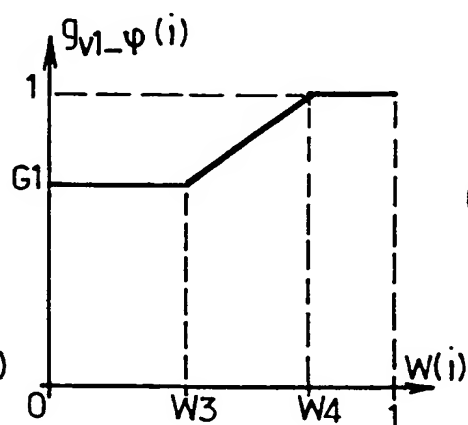
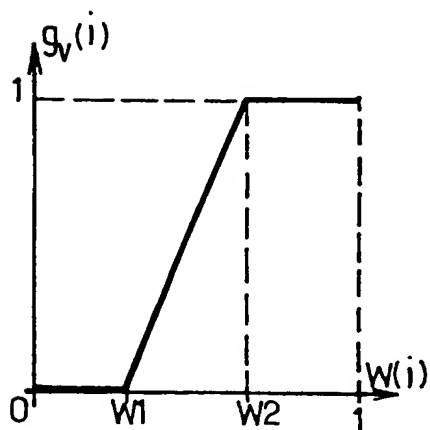
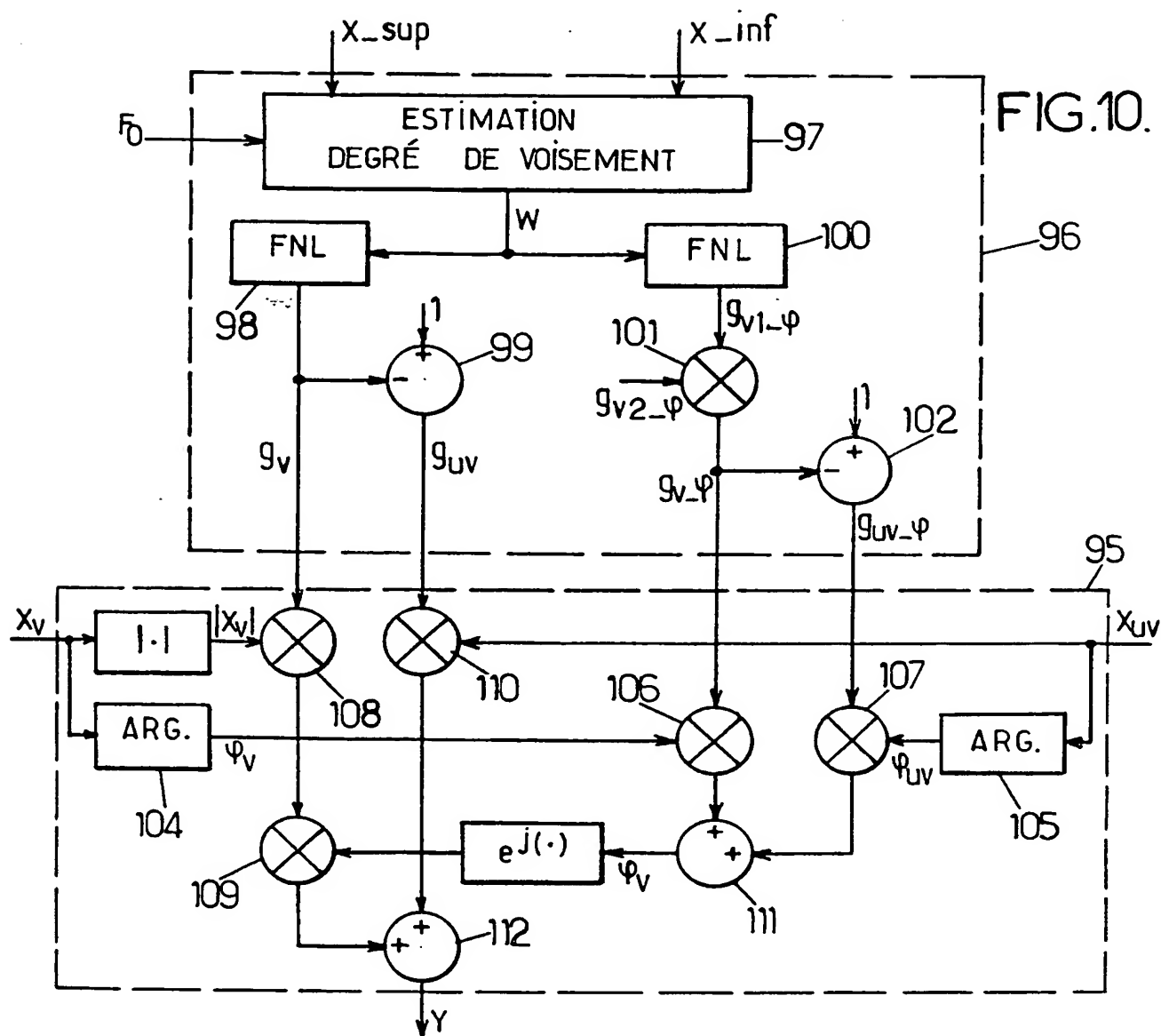
THIS PAGE BLANK (USPTO)

6/13

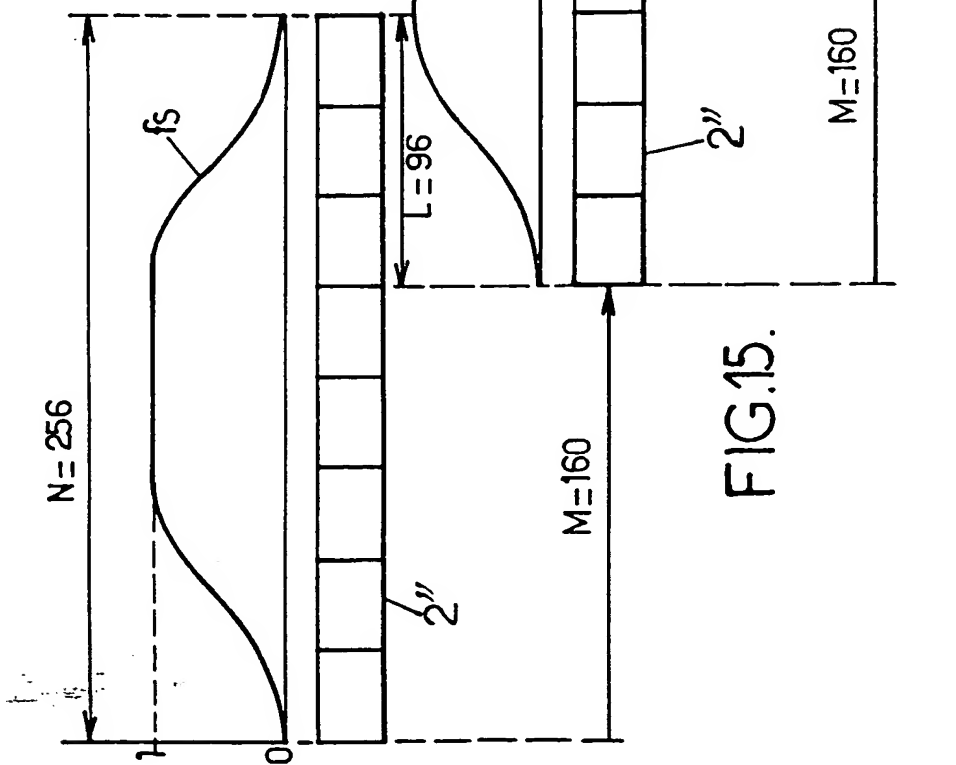
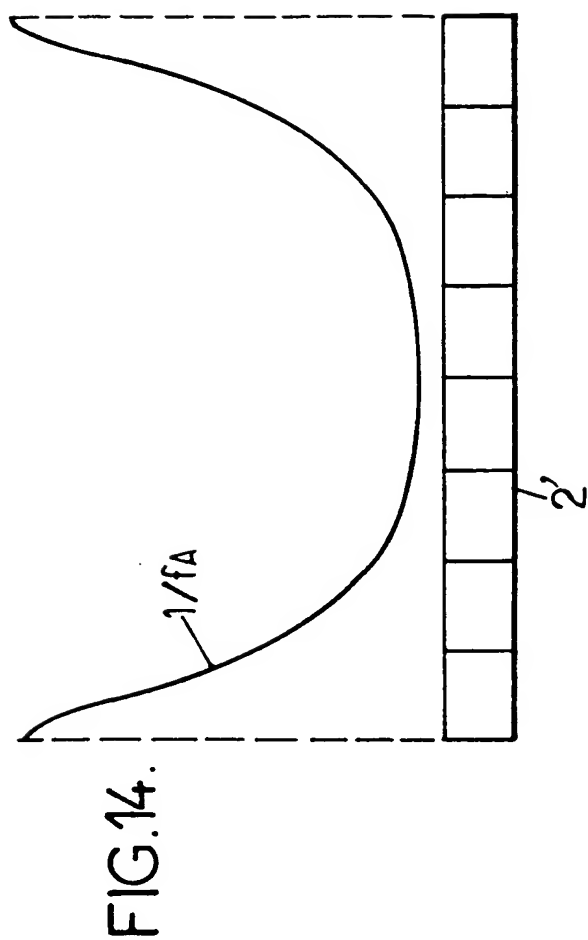


THIS PAGE BLANK (USPTO)

7/13



THIS PAGE BLANK (USPTO)



THIS PAGE BLANK (USPTO

9/13

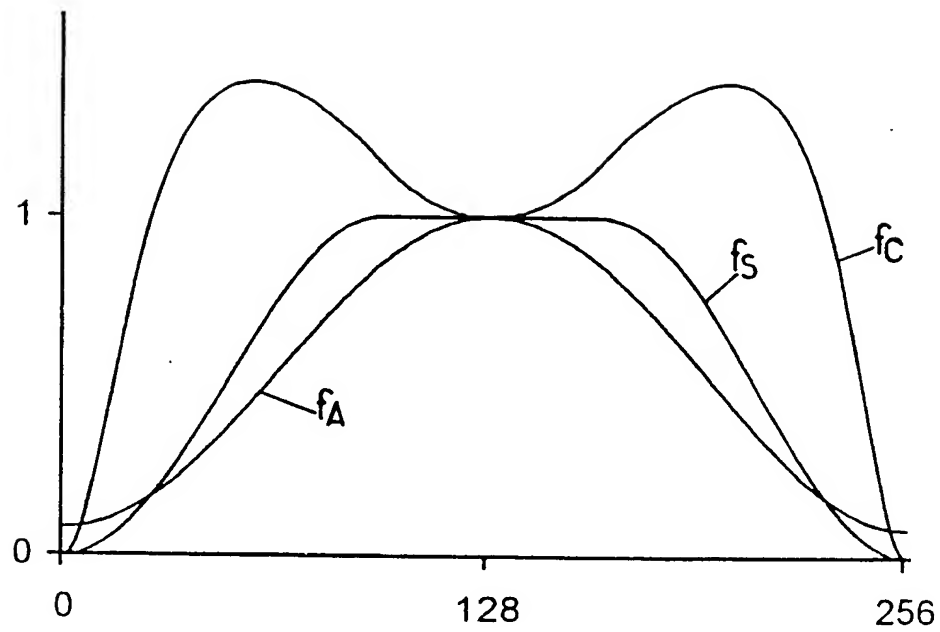


FIG.16.

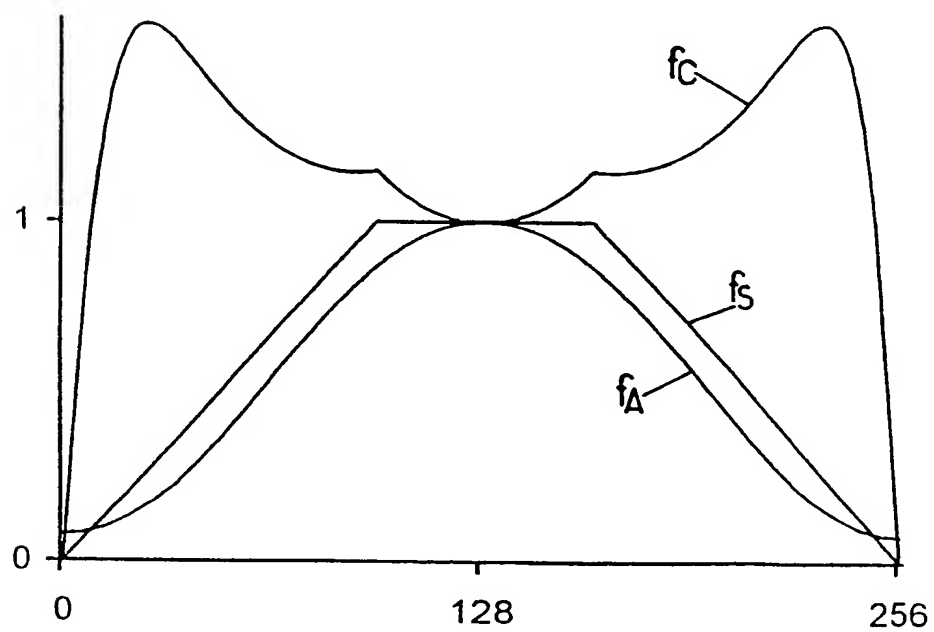
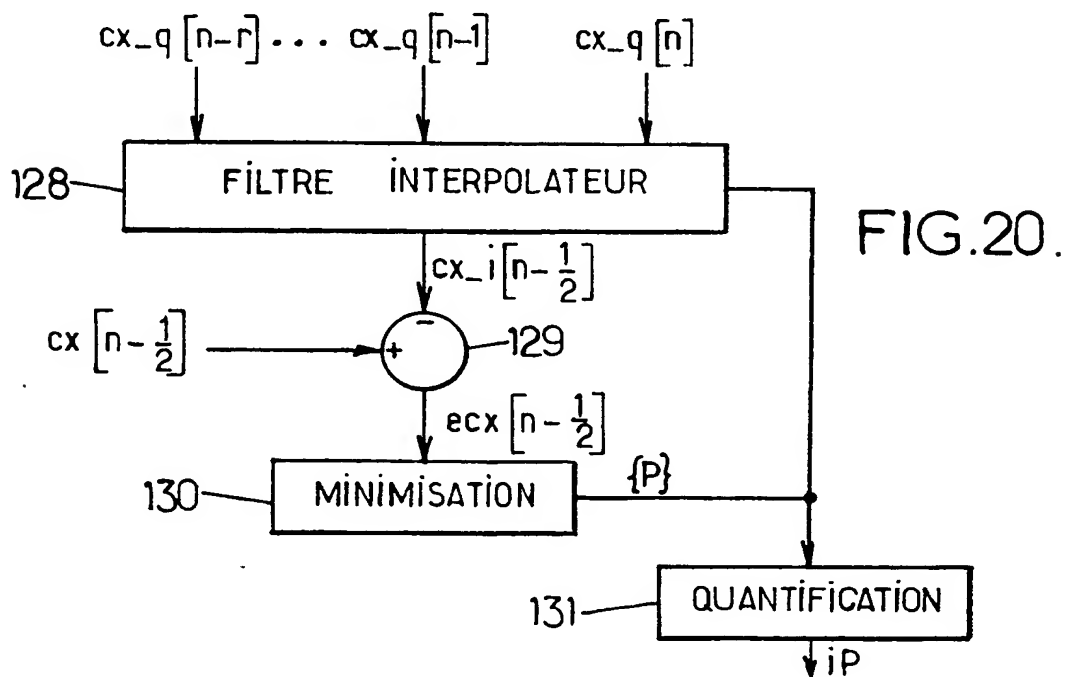
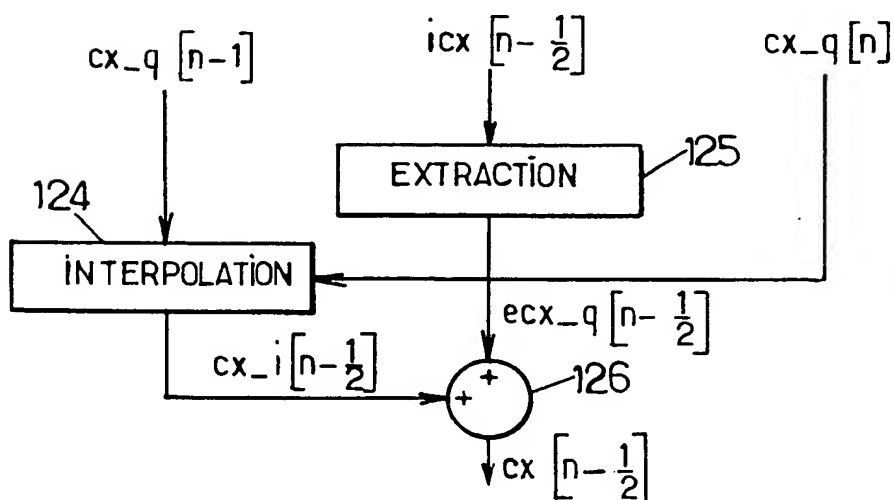
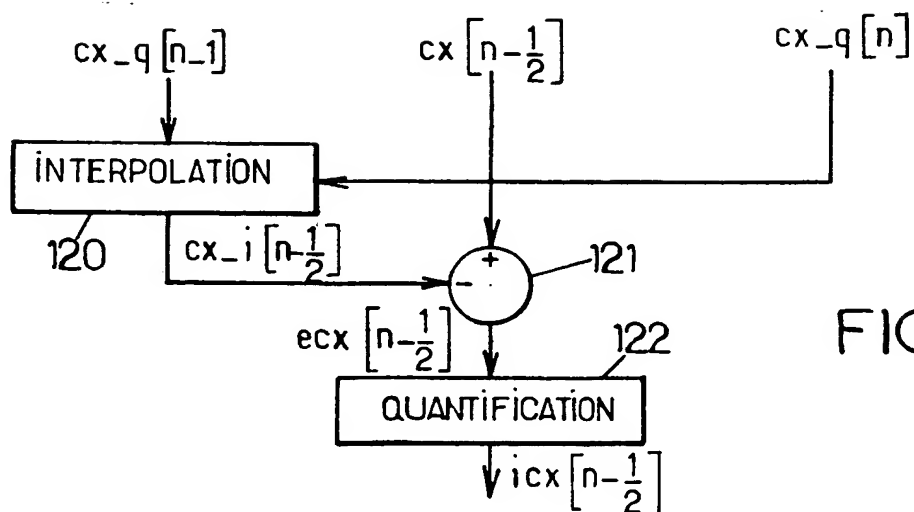


FIG.17.

THIS PAGE BLANK (USPTO)

10/13



THIS PAGE BLANK (USPTO)

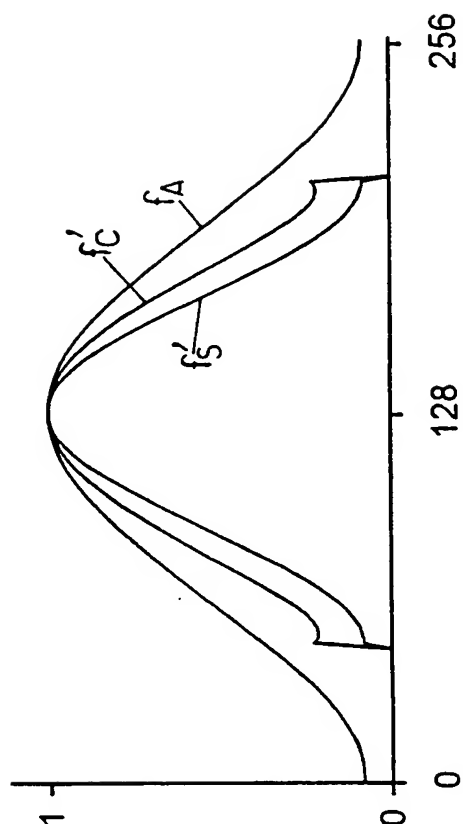


FIG.22.

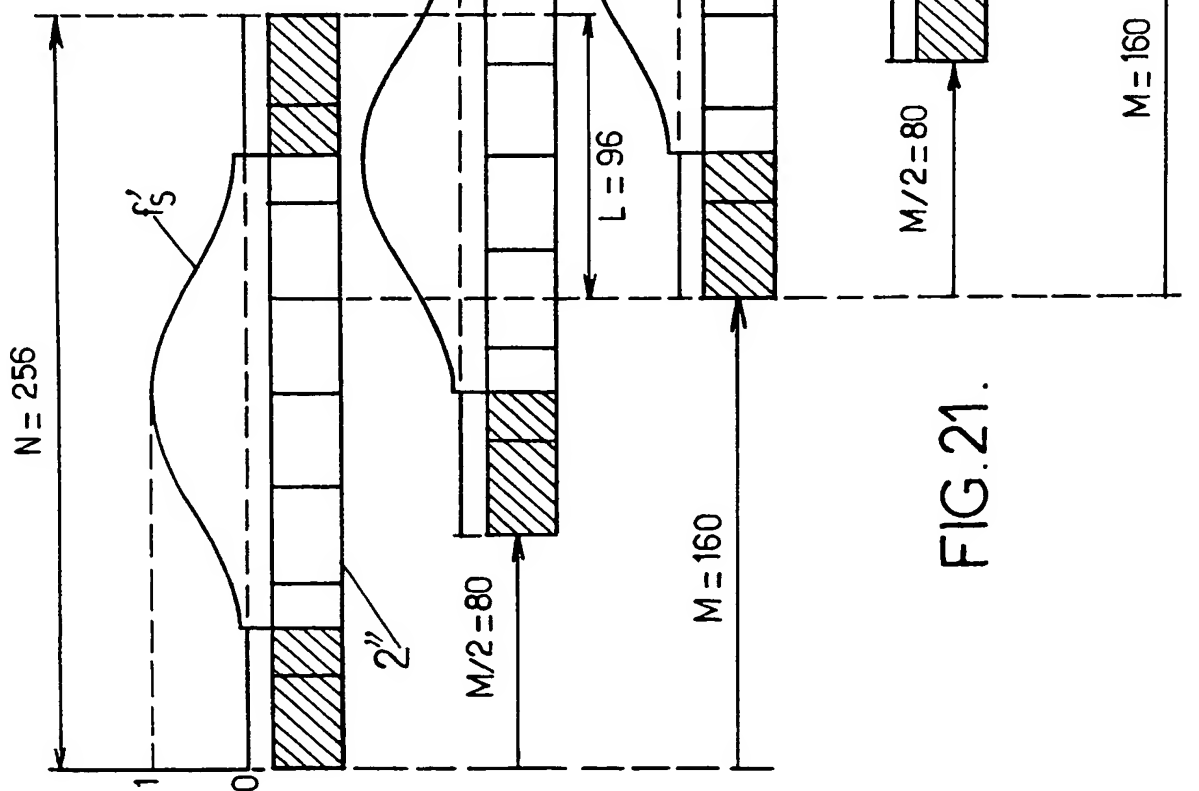


FIG.21.

THIS PAGE BLANK (USPTO

12/13

FIG.23.

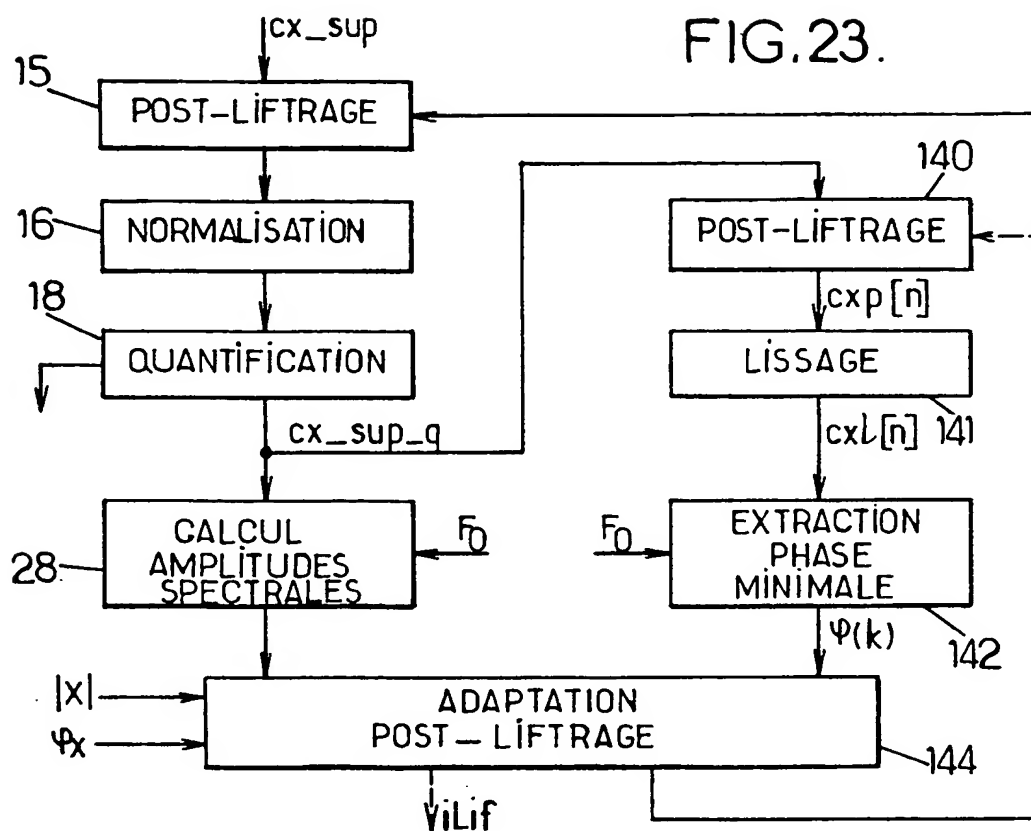
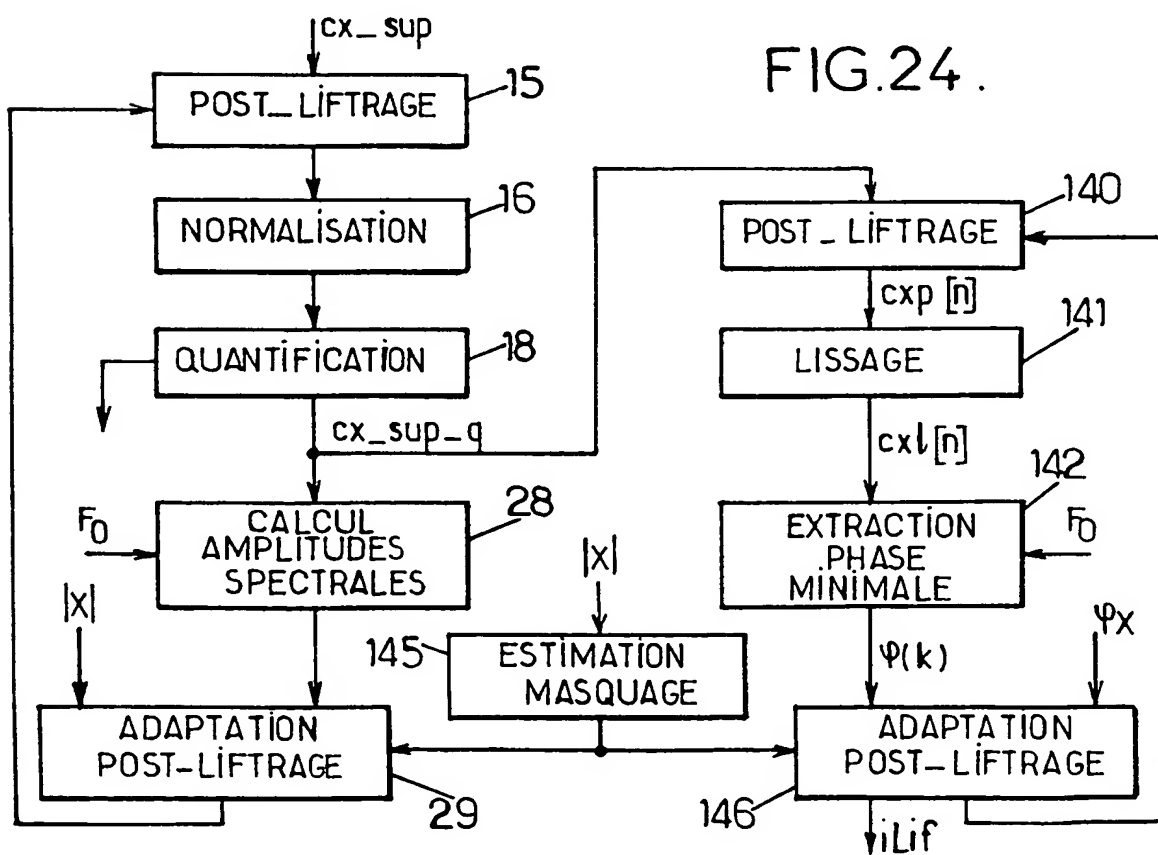


FIG.24.



THIS PAGE BLANK (USPTO)

13/13

FIG. 25.

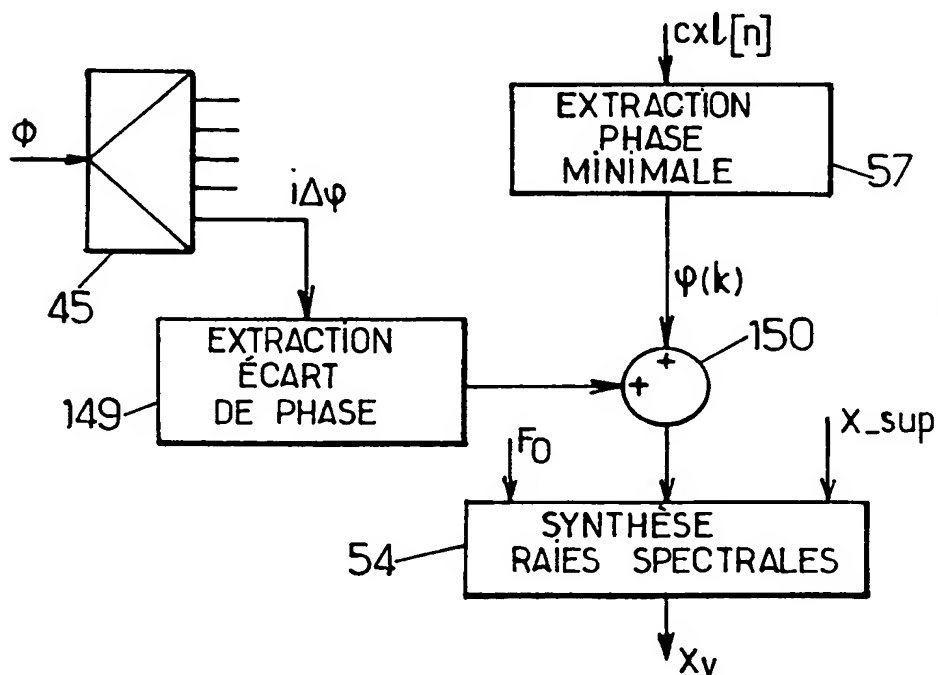
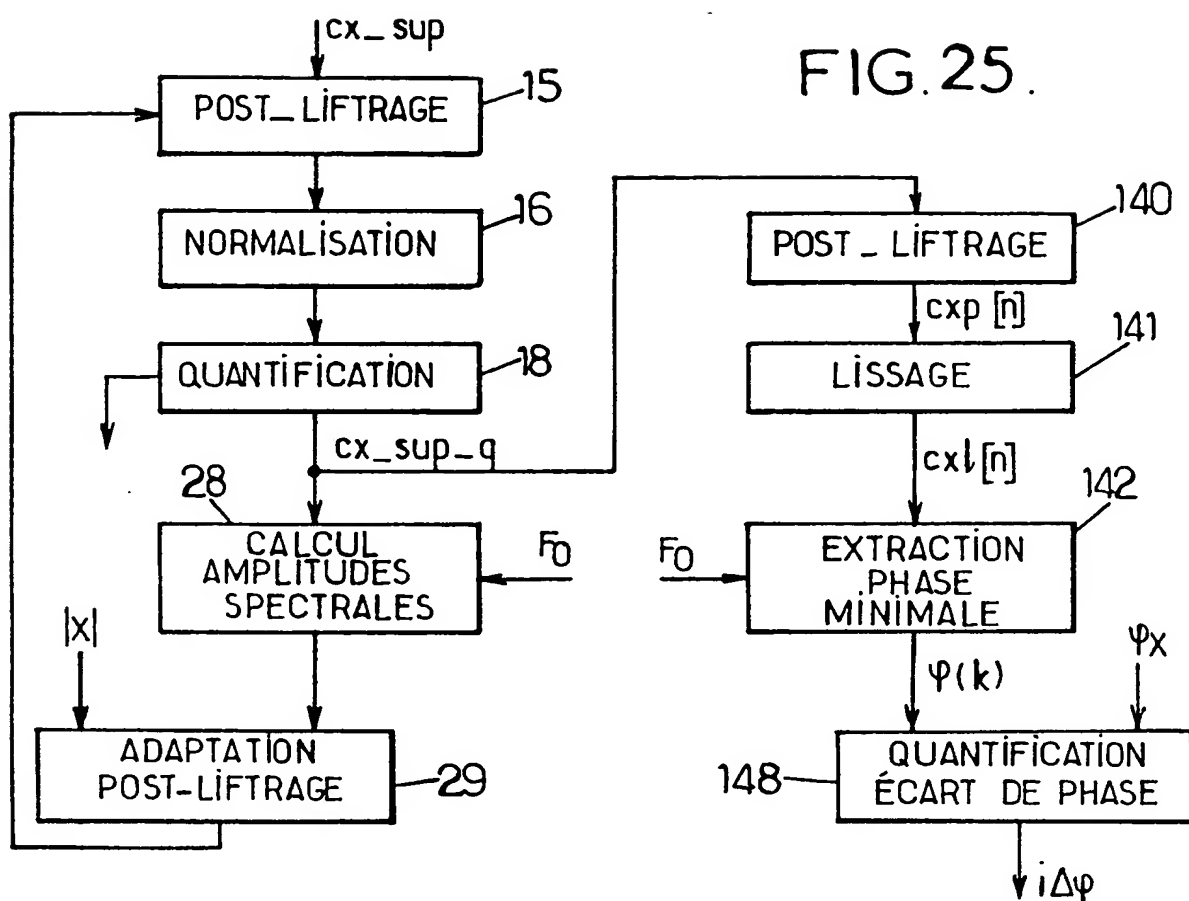


FIG. 26.

THIS PAGE BLANK (USPTO

INTERNATIONAL SEARCH REPORT

Ir. **ational Application No**

PCT/FR 00/01909

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 G10L19/02

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

INSPEC, EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	AHMADI S ET AL: "New techniques for sinusoidal coding of speech at 2400 bps" CONFERENCE RECORD OF THIRTIETH ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS (CAT. NO.96CB36004), PACIFIC GROVE, CA, USA, 3-6 NOV. 1996, pages 770-774 vol.1, XP002138769 1997, Los Alamitos, CA, USA, IEEE Comput. Soc. Press, USA ISBN: 0-8186-7646-9	1,2,16, 17,29
Y	abstract page 770, right-hand column, line 16-38 * pages 772,773 : section "4 Sinusoidal Phase Coding" *	13,26

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *Z* document member of the same patent family

Date of the actual completion of the international search

7 September 2000

Date of mailing of the international search report

14/09/2000

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer _____

Quélavoine, R

INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 00/01909

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 0 709 827 A (MITSUBISHI ELECTRIC CORP) 1 May 1996 (1996-05-01) abstract column 2, line 54 -column 3, line 53 ---	13,26
A	US 5 651 092 A (ISHII ET AL) 22 July 1997 (1997-07-22) claim 1 column 2, line 21-29 column 2, line 39-47 ---	1,4,13, 17,26
A	US 5 701 391 A (PAN ET AL) 23 December 1997 (1997-12-23) abstract column 1, line 62 -column 2, line 15 column 2, line 26-39 column 8, line 15-31 -----	1,2,17

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. l. Application No

PCT/FR 00/01909

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0709827 A	01-05-1996	JP 8123494 A CA 2160749 A CN 1126869 A KR 169020 B US 5724480 A	17-05-1996 29-04-1996 17-07-1996 20-03-1999 03-03-1998
US 5651092 A	22-07-1997	JP 6332496 A CA 2122853 A,C CA 2214585 A DE 69420183 D DE 69420183 T EP 0626674 A EP 0854469 A US 5596675 A	02-12-1994 22-11-1994 22-11-1994 30-09-1999 09-12-1999 30-11-1994 22-07-1998 21-01-1997
US 5701391 A	23-12-1997	AU 7469296 A WO 9716820 A	22-05-1997 09-05-1997

THIS PAGE BLANK (11507)

RAPPORT DE RECHERCHE INTERNATIONALE

De de Internationale No
PCT/FR 00/01909

A. CLASSEMENT DE L'OBJET DE LA DEMANDE
CIB 7 G10L19/02

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)
CIB 7 G10L

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)
INSPEC, EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	AHMADI S ET AL: "New techniques for sinusoidal coding of speech at 2400 bps" CONFERENCE RECORD OF THIRTIETH ASILOMAR CONFERENCE ON SIGNALS, SYSTEMS AND COMPUTERS (CAT. NO.96CB36004), PACIFIC GROVE, CA, USA, 3-6 NOV. 1996, pages 770-774 vol.1, XP002138769 1997, Los Alamitos, CA, USA, IEEE Comput. Soc. Press, USA ISBN: 0-8186-7646-9	1,2,16, 17,29
Y	abrégé page 770, colonne de droite, ligne 16-38 * pages 772,773 : section "4 Sinusoidal Phase Coding" *	13,26

☒ Voir la suite du cadre C pour la fin de la liste des documents

☒ Les documents de familles de brevets sont indiqués en annexe

* Catégories spéciales de documents cités:

"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent

"E" document antérieur, mais publié à la date de dépôt international ou après cette date

"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)

"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens

"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"Z" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

7 septembre 2000

Date d'expédition du présent rapport de recherche internationale

14/09/2000

Nom et adresse postale de l'administration chargée de la recherche internationale
Office Européen des Brevets, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Quélavoine, R

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
Y	EP 0 709 827 A (MITSUBISHI ELECTRIC CORP) 1 mai 1996 (1996-05-01) abrégé colonne 2, ligne 54 -colonne 3, ligne 53	13,26
A	US 5 651 092 A (ISHII ET AL) 22 juillet 1997 (1997-07-22) revendication 1 colonne 2, ligne 21-29 colonne 2, ligne 39-47	1,4,13, 17,26
A	US 5 701 391 A (PAN ET AL) 23 décembre 1997 (1997-12-23) abrégé colonne 1, ligne 62 -colonne 2, ligne 15 colonne 2, ligne 26-39 colonne 8, ligne 15-31	1,2,17

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

De: 1e internationale No

PCT/FR 00/01909

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
EP 0709827 A	01-05-1996	JP 8123494 A	17-05-1996
		CA 2160749 A	29-04-1996
		CN 1126869 A	17-07-1996
		KR 169020 B	20-03-1999
		US 5724480 A	03-03-1998
US 5651092 A	22-07-1997	JP 6332496 A	02-12-1994
		CA 2122853 A,C	22-11-1994
		CA 2214585 A	22-11-1994
		DE 69420183 D	30-09-1999
		DE 69420183 T	09-12-1999
		EP 0626674 A	30-11-1994
		EP 0854469 A	22-07-1998
		US 5596675 A	21-01-1997
US 5701391 A	23-12-1997	AU 7469296 A	22-05-1997
		WO 9716820 A	09-05-1997

THIS PAGE BLANK (USPTO)